# User's Guide II

Descriptive Statistics, Means, Quality Control, and Design of Experiments

**NCSS**
**Statistical System**

**Published by**
**NCSS**
**Dr. Jerry L. Hintze**
**Kaysville, Utah**

# NCSS User's Guide II


**Dr. Jerry L. Hintze**
**Kaysville, Utah 84037**



Direct inquiries to:

NCSS
329 North 1000 East
Kaysville, Utah 84037
Phone (801) 546-0445
Fax (801) 546-3907
Email: support@ncss.com

# NCSS License Agreement

*Important: The enclosed Number Cruncher Statistical System (NCSS) is licensed by Dr. Jerry L. Hintze to customers for their use only on the terms set forth below. Purchasing the system indicates your acceptance of these terms.*

1. **LICENSE.** Dr. Jerry L. Hintze hereby agrees to grant you a non-exclusive license to use the accompanying NCSS program subject to the terms and restrictions set forth in this License Agreement.

2. **COPYRIGHT.** NCSS and its documentation are copyrighted. You may not copy or otherwise reproduce any part of NCSS or its documentation, except that you may load NCSS into a computer as an essential step in executing it on the computer and make backup copies for your use on the same computer.

3. **BACKUP POLICY.** NCSS may be backed up by you for your use on the same machine for which NCSS was purchased.

4. **RESTRICTIONS ON USE AND TRANSFER.** The original and any backup copies of NCSS and its documentation are to be used only in connection with a single computer. You may physically transfer NCSS from one computer to another, provided that NCSS is used in connection with only one computer at a time. You may not transfer NCSS electronically from one computer to another over a network. You may not distribute copies of NCSS or its documentation to others. You may transfer this license together with the original and all backup copies of NCSS and its documentation, provided that the transferee agrees to be bound by the terms of this License Agreement. Neither NCSS nor its documentation may be modified or translated without written permission from Dr. Jerry L. Hintze.

*You may not use, copy, modify, or transfer **NCSS**, or any copy, modification, or merged portion, in whole or in part, except as expressly provided for in this license.*

5. **NO WARRANTY OF PERFORMANCE.** Dr. Jerry L. Hintze does not and cannot warrant the performance or results that may be obtained by using NCSS. Accordingly, NCSS and its documentation are licensed "as is" without warranty as to their performance, merchantability, or fitness for any particular purpose. The entire risk as to the results and performance of NCSS is assumed by you. Should NCSS prove defective, you (and not Dr. Jerry L. Hintze nor his dealers) assume the entire cost of all necessary servicing, repair, or correction.

6. **LIMITED WARRANTY ON CD.** To the original licensee only, Dr. Jerry L. Hintze warrants the medium on which NCSS is recorded to be free from defects in materials and faulty workmanship under normal use and service for a period of ninety days from the date NCSS is delivered. If, during this ninety-day period, a defect in a CD should occur, the CD may be returned to Dr. Jerry L. Hintze at his address, or to the dealer from which NCSS was purchased, and NCSS will replace the CD without charge to you, provided that you have sent a copy of your receipt for NCSS. Your sole and exclusive remedy in the event of a defect is expressly limited to the replacement of the CD as provided above.

Any implied warranties of merchantability and fitness for a particular purpose are limited in duration to a period of ninety (90) days from the date of delivery. If the failure of a CD has resulted from accident, abuse, or misapplication of the CD, Dr. Jerry L. Hintze shall have no responsibility to replace the CD under the terms of this limited warranty. This limited warranty gives you specific legal rights, and you may also have other rights, which vary from state to state.

7. **LIMITATION OF LIABILITY.** Neither Dr. Jerry L. Hintze nor anyone else who has been involved in the creation, production, or delivery of NCSS shall be liable for any direct, incidental, or consequential damages, such as, but not limited to, loss of anticipated profits or benefits, resulting from the use of NCSS or arising out of any breach of any warranty. Some states do not allow the exclusion or limitation of direct, incidental, or consequential damages, so the above limitation may not apply to you.

8. **TERM.** The license is effective until terminated. You may terminate it at any time by destroying NCSS and documentation together with all copies, modifications, and merged portions in any form. It will also terminate if you fail to comply with any term or condition of this License Agreement. You agree upon such termination to destroy NCSS and documentation together with all copies, modifications, and merged portions in any form.

9. **YOUR USE OF NCSS ACKNOWLEDGES** that you have read this customer license agreement and agree to its terms. You further agree that the license agreement is the complete and exclusive statement of the agreement between us and supersedes any proposal or prior agreement, oral or written, and any other communications between us relating to the subject matter of this agreement.

Dr. Jerry L. Hintze & **NCSS**, Kaysville, Utah

# Preface

Number Cruncher Statistical System (**NCSS**) is an advanced, easy-to-use statistical analysis software package. The system was designed and written by Dr. Jerry L. Hintze over the last several years. Dr. Hintze drew upon his experience both in teaching statistics at the university level and in various types of statistical consulting.

The present version, written for 32-bit versions of Microsoft Windows (95, 98, ME, 2000, NT, etc.) computer systems, is the result of several iterations. Experience over the years with several different types of users has helped the program evolve into its present form.

Statistics is a broad, rapidly developing field. Updates and additions are constantly being made to the program. If you would like to be kept informed about updates, additions, and corrections, please send your name, address, and phone number to:

User Registration
NCSS
329 North 1000 East
Kaysville, Utah 84037

or Email you name, address, and phone number to:

Sales@NCSS.COM

**NCSS** maintains a website at **WWW.NCSS.COM** where we make the latest edition of NCSS available for free downloading. The software is password protected, so only users with valid serial numbers may use this downloaded edition. We hope that you will download the latest edition routinely and thus avoid any bugs that have been corrected since you purchased your copy.

**NCSS** maintains the following program and documentation copying policy. Copies are limited to a one person / one machine basis for "BACKUP" purposes only. You may make as many backup copies as you wish. Further distribution constitutes a violation of this copy agreement and will be prosecuted to the fullest extent of the law.

**NCSS** is not "copy protected." You may freely load the program onto your hard disk. We have avoided copy protection in order to make the system more convenient for you. Please honor our good faith (and low price) by avoiding the temptation to distribute copies to friends and associates.

We believe this to be an accurate, exciting, easy-to-use system. If you find any portion that you feel needs to be changed, please let us know. Also, we openly welcome suggestions for additions to the system.

# User's Guide II
## Table of Contents

# User's Guide I
## Table of Contents

# User's Guide III
## Table of Contents

# User's Guide IV
## Table of Contents

# User's Guide V
## Table of Contents

Chapter 200

# Descriptive Statistics

## Introduction

This procedure summarizes variables both statistically and graphically. Information about the location (center), spread (variability), and distribution is provided. The procedure provides a large variety of statistical information about a single variable.

## Kinds of Research Questions

The use of this module for a single variable is generally appropriate for one of four purposes: numerical summary, data screening, outlier identification (which sometimes is incorporated into data screening), and distributional shape. We will briefly discuss each of these now.

## Numerical Descriptors

The numerical descriptors of a sample are called statistics. These statistics may be categorized as location, spread, shape indicators, percentiles, and interval estimates.

### Location or Central Tendency

One of the first impressions that we like to get from a variable is its general location. You might think of this as the center of the variable on the number line. The average (mean) is a common measure of location. When investigating the center of a variable, the main descriptors are the mean, median, mode, and the trimmed mean. Other averages, such as the geometric and harmonic mean, have specialized uses. We will now briefly compare these measures.

If the data come from the normal distribution, the mean, median, mode, and the trimmed mean are all equal. If the mean and median are very different, most likely there are outliers in the data or the distribution is skewed. If this is the case, the median is probably a better measure of location. The mean is very sensitive to extreme values and can be seriously contaminated by just one observation.

A compromise between the mean and median is given by the trimmed mean (where a predetermined number of observations are trimmed from each end of the data distribution). This trimmed mean is more robust than the mean but more sensitive than the median. Comparison of the trimmed mean to the median should show the trimmed mean approaching the median as the

degree of trimming increases. If the trimmed mean converges to the median for a small degree of trimming, say 5 or 10%, the number of outliers is relatively few.

## Variability, Dispersion, or Spread

After establishing the center of a variable's values, the next question is how closely the data fall about this center. The pattern of the values around the center is called the *spread*, *dispersion*, or *variability*. There are numerous measures of variability: range, variance, standard deviation, interquartile range, and so on. All of these measures of dispersion are affected by outliers to some degree, but some do much better than others.

The *standard deviation* is one of the most popular measures of dispersion. Unfortunately, it is greatly influenced by outlying observations and by the overall shape of the distribution. Because of this, various substitutes for it have been developed. It will be up to you to decide which is best in a given situation.

## Shape

The shape of the distribution describes the pattern of the values along the number line. Are there a few unique values that occur over and over, or is there a continuum? Is the pattern symmetric or asymmetric? Are the data bell shaped? Do they seem to have a single center or are there several areas of clumping? These are all aspects of the shape of the distribution of the data.

Two of the most popular measures of shape are skewness and kurtosis. *Skewness* measures the direction and lack of *symmetry*. The more skewed a distribution is, the greater the need for using robust estimators, such as the median and the interquartile range. Positive skewness indicates a longtailedness to the right while negative skewness indicates longtailedness to the left. *Kurtosis* measures the heaviness of the tails. A kurtosis value less than three indicates lighter tails than a normal distribution. Kurtosis values greater than three indicate heavier tails than a normal distribution.

The measures of shape require more data to be accurate. For example, a reasonable estimate of the mean may require only ten observations in a random sample. The standard deviation will require at least thirty. A reasonably detailed estimate of the shape (especially if the tails are important) will require several hundred observations.

## Percentiles

Percentiles are extremely useful for certain applications as well as for cases when the distribution is very skewed or contaminated by outliers. If the distribution of the variable is skewed, you might want to use the exact interval estimates for the percentiles.

## Confidence Limits or Interval Estimates

An interval estimate of a statistic gives a range of its possible values. Confidence limits are a special type of interval estimate that have, under certain conditions, a level of confidence or probability attached to them.

If the assumption of normality is valid, the confidence intervals for the mean, variance, and standard deviation are valid. However, the standard error of each of these intervals depends on the sample standard deviation and the sample size. If the sample standard deviation is inaccurate, these other measures will be also. The bottom line is that outliers not only affect the standard

deviation but also all confidence limits that use the sample standard deviation. It should be obvious then that the standard deviation is a critical measure of dispersion in parametric methods.

# Data Screening

Data screening involves missing data, data validity, and outliers. If these issues are not dealt with prior to the use of descriptive statistics, errors in interpretations are very likely.

## Missing Data

Whenever data are missing, questions need to be asked.

1. Is the missingness due to incomplete data collection? If so, try to complete the data collection.
2. Is the missingness due to nonresponse from a survey? If so, attempt to collect data from the nonresponders.
3. Are the missing data due to a censoring of data beyond or below certain values? If so, some different statistical tools will be needed.
4. Is the pattern of missingness random? If only a few data points are missing from a large data set and the pattern of missingness is random, there is little to be concerned with. However, if the data set is small or moderate in size, any degree of missingness could cause bias in interpretations.

Whenever missing values occur without answers to the above questions, there is little that can be done. If the distributional shape of the variable is known and there are missing data for certain percentiles, estimates could be made for the missing values. If there are other variables in the data set as well and the pattern of missingness is random, multiple regression and multivariate methods can be used to estimate the missing values.

## Data Validity

Data validity needs to be confirmed prior to any statistical analysis, but it usually begins after a univariate descriptive analysis. Extremes or outliers for a variable could be due to a data entry error, to an incorrect or inappropriate specification of a missing code, to sampling from a population other than the intended one, or due to a natural abnormality that exists in this variable from time to time. The first two cases of invalid data are easily corrected. The latter two require information about the distribution form and necessitate the use of regression or multivariate methods to re-estimate the values.

## Outliers

Outliers in a univariate data set are defined as observations that appear to be inconsistent with the rest of the data. An outlier is an observation that sticks out at either end of the data set.

The visualization of univariate outliers can be done in three ways: with the stem-and-leaf plot, with the box plot, and with the normal probability plot. In each of these informal methods, the outlier is far removed from the rest of the data. A word of caution: the box plot and the normal probability plot evaluate the potentiality of an outlier assuming the data are normally distributed. If the variable is not normally distributed, these plots may indicate many outliers. You must be

careful about checking what distributional assumptions are behind the outliers you may be looking for.

Outliers can completely distort descriptive statistics. For instance, if one suspects outliers, a comparison of the mean, median, mode, and trimmed mean should be made. If the outliers are only to one side of the mean, the median is a better measure of location. On the other hand, if the outliers are equally divergent on each side of the center, the mean and median will be close together, but the standard deviation will be inflated. The interquartile range is the only measure of variation not greatly affected by outliers. Outliers may also contaminate measures of skewness and kurtosis as well as confidence limits.

This discussion has focused on univariate outliers, in a simplistic way. If the data set has several variables, multiple regression and multivariate methods must be used to identify these outliers.

## Normality

A primary use of descriptive statistics is to determine whether the data are normally distributed. If the variable is normally distributed, you can use parametric statistics that are based on this assumption. If the variable is not normally distributed, you might try a transformation on the variable (such as, the natural log or square root) to make the data normal. If a transformation is not a viable alternative, nonparametric methods that do not require normality should be used.

***NCSS*** provides seven tests to formally test for normality. If a variable fails a normality test, it is critical to look at the box plot and the normal probability plot to see if an outlier or a small subset of outliers has caused the nonnormality. A pragmatic approach is to omit the outliers and rerun the tests to see if the variable now passes the normality tests.

Always remember that a reasonably large sample size is necessary to detect normality. Only extreme types of nonnormality can be detected with samples less than fifty observations.

There is a common misconception that a histogram is always a valid graphical tool for assessing normality. Since there are many subjective choices that must be made in constructing a histogram, and since histograms generally need large sample sizes to display an accurate picture of normality, preference should be given to other graphical displays such as the box plot, the density trace, and the normal probability plot.

# Data Structure

The data are contained in a single variable.

**SAMPLE dataset (subset)**

| Height |
|---|
| 64 |
| 63 |
| 67 |
| . |
| . |
| . |

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

Following is a list of the procedure's options.

## Variables Tab

The options on this panel specify which variables to use.

### Data Variables

#### Variable(s)

Specify a list of one or more variables upon which the univariate statistics are to be generated. You can double-click the field or single click the button on the right of the field to bring up the Variable Selection window.

### Grouping Variables

#### Group (1-5) Variable

You can select up to five categorical variables. When one or more of these are specified, a separate set of reports is generated for each unique set of values for these variables.

### Frequency Variable

#### Frequency Variable

This optional variable specifies the number of observations that each row represents. When omitted, each row represents a single observation. If your data is the result of a previous summarization, you may want certain rows to represent several observations. Note that negative values are treated as a zero weight and are omitted. This is one way of weighting your data.

### Data Transformation Options

#### Exponent

Occasionally, you might want to obtain a statistical report on the square root or square of your variable. This option lets you specify an on-the-fly transformation of the variable. The form of this transformation is $X = Y^A$, where Y is the original value, A is the selected exponent, and X is the value that is summarized.

#### Additive Constant

Occasionally, you might want to obtain a statistical report on a transformed version of a variable. This option lets you specify an on-the-fly transformation of the variable. The form of this transformation is X = Y+B, where Y is the original value, B is the selected value, and X is the value that is summarized.

Note that if you apply both the *Exponent* and the *Additive Constant*, the form of the transformation is $X = (Y+B)^A$.

# Reports Tab

The options on this panel control the format of the report.

## Select Reports

### Summary Section … Percentile Section

Each of these options indicates whether to display the indicated report.

## Select Plots

### Stem Leaf, Histogram, Probability Plot

Each of these options indicates whether to display the indicated plot.

## Report Options

### Alpha Level

The value of alpha for the confidence limits and rejection decisions. Usually, this number will range from 0.1 to 0.001. The default value of 0.05 results in 95% confidence limits.

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Value Labels

This option applies to the *Group Variable(s)*. It lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

## Report Options - Decimal Places

### Values, Means, Probabilities

Specify the number of decimal places when displaying this item. Select ‘General’ to display all possible decimal places.

## Report Options - Percentiles

### Percentile Type

This selects from five methods used to calculate the $p^{th}$ percentile, $z_p$. The first option, *Xp(n+1)*, gives the common value of the median. These options are:

- **AveXp(n+1)**

  The $100p^{th}$ percentile is computed as

  $$Z_p = (1\text{-}g)X_{[k1]} + gX_{[k2]}$$

  where $k1$ equals the integer part of $p(n+1)$, $k2=k1+1$, $g$ is the fractional part of $p(n+1)$, and $X_{[k]}$ is the $k^{th}$ observation when the data are sorted from lowest to highest.

- **AveXp(n)**

  The $100p^{th}$ percentile is computed as

  $$Z_p = (1\text{-}g)X_{[k1]} + gX_{[k2]}$$

  where $k1$ equals the integer part of $np$, $k2=k1+1$, $g$ is the fractional part of $np$, and $X_{[k]}$ is the $k^{th}$ observation when the data are sorted from lowest to highest.

- **Closest to np**

  The $100p^{th}$ percentile is computed as

  $$Z_p = X_{[k1]}$$

  where $k1$ equals the integer that is closest to $np$ and $X_{[k]}$ is the $k^{th}$ observation when the data are sorted from lowest to highest.

- **EDF**

  The $100p^{th}$ percentile is computed as

  $$Z_p = X_{[k1]}$$

  where $k1$ equals the integer part of $np$ if $np$ is exactly an integer or the integer part of $np+1$ if $np$ is not exactly an integer. $X_{[k]}$ is the $k^{th}$ observation when the data are sorted from lowest to highest. Note that EDF stands for empirical distribution function.

- **EDF w/Ave**

  The $100p^{th}$ percentile is computed as

  $$Z_p = (X_{[k1]} + X_{[k2]})/2$$

  where $k1$ and $k2$ are defined as follows: If $np$ is an integer, $k1=k2=np$. If $np$ is not exactly an integer, $k1$ equals the integer part of $np$ and $k2 = k1+1$. $X_{[k]}$ is the $k^{th}$ observation when the data are sorted from lowest to highest. Note that EDF stands for empirical distribution function.

- **Smallest Percentile**

  By default, the smallest percentile displayed is the 1st percentile. This option lets you change this value to any value between 0 and 100. For example, you might enter 2.5 to see the $2.5^{th}$ percentile.

- **Largest Percentile**

  By default, the largest percentile displayed is the 99th percentile. This option lets you change this value to any value between 0 and 100. For example, you might enter 97.5 to see the $97.5^{th}$ percentile.

# Probability Plot Tab

The options on this panel control the appearance of the probability plot.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Probability Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Probability Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

# Histogram Tab

The options on this panel control the appearance of the histogram.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Histogram Settings

### Plot Style File

Designate a histogram style file. This file sets all histogram options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Histogram procedure.

### Number of Bars

Specify the number of intervals, bins, or bars used in the histogram.

## Titles

### Plot Title

This is the text of the title. The characters *{X}* are replaced by the name of the variable. Press the button on the right of the field to specify the font of the text.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running Descriptive Statistics

This section presents a detailed example of how to run a descriptive statistics report on the *Height* variable in the SAMPLE database. To run this example, take the following steps (note that step 1 is not necessary if the SAMPLE dataset is open):

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Descriptive Statistics window.

**1 Open the SAMPLE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Sample.s0**.
- Click **Open**.

**2 Open the Descriptive Statistics window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics**. The Descriptive Statistics procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the Height variable.**

- On the Descriptive Statistics window, select the **Variables tab**. (This is the default.)
- Double-click in the **Variables** text box. This will bring up the variable selection window.
- Select **Height** from the list of variables and then click **Ok**. The word "Height" will appear in the Variables box. Remember that you could have entered a "**1**" here signifying the first (left-most) variable on the dataset.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports and charts will be displayed in the Output window.

## Descriptive Statistics Report

This report is rather large and complicated, so we will define each section separately. Usually, you will focus on only a few items from this report. Unfortunately, each user wants a different few items, so we had to include much more than any one user needs!

Several of the formulas involve both raw and central moments. The raw moments are defined as:

$$m'_r = \frac{\sum_{i=1}^{n} x_i^r}{n}$$

The central moments are defined as:

$$m_r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^r}{n}$$

Large sample estimates of the standard errors are provided for several statistics. These are based on the following formula from Kendall and Stuart (1987):

$$Var(m_r) = \frac{m_{2r} - m_r^2 + 4m_2 m_{r-1}^2 - 2r m_{r-1} m_{r+1}}{n}$$

$$Var(g(x)) = \left(\frac{dg}{dx}\right)^2 Var(x)$$

## Summary Section

**Summary Section of Height**

| Count | Mean | Standard Deviation | Standard Error | Minimum | Maximum | Range |
|---|---|---|---|---|---|---|
| 20 | 62.1 | 8.441128 | 1.887493 | 51 | 79 | 28 |

### Count

This is the number of nonmissing values. If no frequency variable was specified, this is the number of nonmissing rows.

### Mean

This is the average of the data values. (See *Means Section* below.)

### Standard Deviation

This is the standard deviation of the data values. (See *Variation Section* below.)

### Standard Error

This is the standard error of the mean. (See *Means Section* below.)

### Minimum

The smallest value in this variable.

### Maximum

The largest value in this variable.

### Range

The difference between the largest and smallest values for a variable. If the data for a given variable is normally distributed, a quick estimate of the standard deviation can be made by dividing the range by six.

## Count Section

**Counts Section of Height**

| Rows | Sum of Frequencies | Missing Values | Distinct Values | Sum | Total Sum Squares | Adjusted Sum Squares |
|---|---|---|---|---|---|---|
| 75 | 20 | 55 | 14 | 1242 | 78482 | 1353.8 |

### Rows

This is the total number of rows available in this variable.

### Sum of Frequencies

This is the number of nonmissing values. If no frequency variable was specified, this is the number of nonmissing rows.

### Missing Values

The number of missing (empty) rows.

### Distinct Values

This is the number of unique values in this variable. This value is useful for finding data entry errors and for determining if a variable is continuous or discrete.

### Sum

This is the sum of the data values.

### Total Sum Squares

This is the sum of the squared values of the variable. It is sometimes referred to as the *unadjusted sum of squares.* It is reported for its usefulness in calculating other statistics and is not interpreted directly.

$$sum\ squares = \sum_{i=1}^{n} x_i^2$$

### Adjusted Sum Squares

This is the sum of the squared differences from the mean.

$$sum\ squares = \sum_{i=1}^{n}(x_i - \bar{x})^2$$

## Means Section

**Means Section of Height**

| Parameter | Mean | Median | Geometric Mean | Harmonic Mean | Sum | Mode |
|---|---|---|---|---|---|---|
| Value | 62.1 | 59.5 | 61.57052 | 61.05865 | 1242 | 52 |
| Std Error | 1.887493 | | | | 37.74987 | |
| 95% LCL | 58.14943 | 56 | | | 1162.989 | |
| 95% UCL | 66.05057 | 67 | | | 1321.011 | |
| T-Value | 32.9008 | | | | | |
| Prob Level | .000000 | | | | | |
| Count | 20 | | 20 | 20 | | 3 |

### Mean

This is the average of the data values.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

### Std Error (Mean)

This is the standard error of the mean. This is the estimated standard deviation for the distribution of sample means for an infinite population.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

### LCL and 95% UCL of the Mean

This is the upper and lower values of a 100(1-α) interval estimate for the mean based on a t distribution with *n-1* degrees of freedom. This interval estimate assumes that the population standard deviation is not known and that the data for this variable are normally distributed.

$$\bar{x} \pm t_{a/2,n-1} s_{\bar{x}}$$

### T-Value (Mean)

This is the t-test value for testing that the sample mean is equal to zero versus the alternative that it is not. The degrees of freedom for this t-test are *n-1*. The variable that is being tested must be approximately normally distributed for this test to be valid.

$$t_{\alpha/2,n-1} = \frac{\bar{x}}{s_{\bar{x}}}$$

### Prob Level (Mean)

This is the significance level of the above t-test, assuming a two-tailed test. Generally, this p-value is compared to the level of significance, .05 or .01, chosen by the researcher. If the p-value is less than the pre-determined level of significance, the sample mean is different from zero.

### Median

The value of the median. The median is the 50th percentile of the data set. It is the point that splits the data base in half. The value of the percentile depends upon the percentile method that was selected.

### LCL and 95% UCL of the Median

These are the values of an exact confidence interval for the median. These exact confidence intervals are discussed in the *Percentile Section*.

### Geometric Mean

The geometric mean (GM) is an alternative type of mean that is used for business, economic, and biological applications. Only nonnegative values are used in the computation. If one of the values is zero, the geometric mean is defined to be zero.

One example of when the GM is appropriate is when a variable is the product of many small effects combined by multiplication instead of addition.

$$GM = \left( \prod_{i=1}^{n} x_i \right)^{1/n}$$

An alternative form, showing the GM's relationship to the arithmetic mean, is:

$$GM = \exp\left( \frac{1}{n} \sum \ln(x_i) \right)$$

### Count for Geometric Mean

The number of positive numbers used in computing the geometric mean.

### Harmonic Mean

The harmonic mean is used to average rates. For example, suppose we want the average speed of a bus that travels a fixed distance every day at speeds $s_1$, $s_2$, and $s_3$. The average speed, found by dividing the total distance by the total time, is equal to the harmonic mean of the three speeds. The harmonic mean is appropriate when the distance is constant from trial to trial and the time required was variable. However, if the times were constant and the distances were variable, the arithmetic mean would have been appropriate.

Only nonzero values may be used in its calculation.

$$HM = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

### Count for the Harmonic Mean

The number of nonzero numbers used in computing the harmonic mean.

### Sum

This is the sum of the data values. The standard error and confidence limits are found by multiplying the corresponding values for the mean by the sample size, *n*.

### Std Error of Sum

This is the standard deviation of the distribution of sums. With this standard error, confidence intervals and hypothesis testing can be done for the sum. The assumptions for the interval estimate of the mean must also hold here.

$$s_{sum} = ns_{\bar{x}}$$

### Mode

This is the most frequently occurring value in a data.

### Mode Count

This is a count of the most frequently occurring value, i.e., frequency.

## Variation Section

**Variation Section of Height**

| Parameter | Variance | Standard Deviation | Unbiased Std Dev | Std Error of Mean | Interquartile Range | Range |
|---|---|---|---|---|---|---|
| Value | 71.25263 | 8.441128 | 8.552877 | 1.887493 | 14 | 28 |
| Std Error | 17.01612 | 1.425427 | | 0.3187352 | | |
| 95% LCL | 41.20865 | 6.419396 | | 1.435421 | | |
| 95% UCL | 152.0011 | 12.32887 | | 2.756819 | | |

### Variance

The sample variance, $s^2$, is a popular measure of dispersion. It is an average of the squared deviations from the mean.

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

### Std Error of Variance

This is a large sample estimate of the standard error of $s^2$ for an infinite population.

### LCL of the Variance

This is the lower value of a 100(1-α) interval estimate for the variance based on the chi-squared distribution with *n-1* degrees of freedom. This interval estimate assumes that the variable is normally distributed.

$$LCL = \frac{s^2(n-1)}{\chi^2_{\alpha/2,n-1}}$$

### UCL of the Variance

This is the upper value of a 100(1-α) interval estimate for the variance based on the chi-squared distribution with *n-1* degrees of freedom. This interval estimate assumes that the variable is normally distributed.

$$UCL = \frac{s^2(n-1)}{\chi^2_{1-\alpha/2,n-1}}$$

### Standard Deviation

The sample standard deviation, *s*, is a popular measure of dispersion. It measures the average distance between a single observation and its mean. The use of *n-1* in the denominator instead of the more natural *n* is often of concern. It turns out that if *n* (instead of *n-1*) were used, a biased estimate of the population standard deviation would result. The use of *n-1* corrects for this bias.

Unfortunately, *s* is inordinately influenced by outliers. For this reason, you must always check for outliers in your data before you use this statistic. Also, *s* is a biased estimator of the population standard deviation. An unbiased estimate, calculated by adjusting *s*, is given under the heading *Unbiased Std Dev*.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Another form of the above formula that shows that the standard deviation is proportional to the difference between each pair of observations. Notice that the sample mean does not enter into this second formulation.

$$s = \sqrt{\frac{\sum_{all\ i,j\ where\ i<j}(x_i - x_j)^2}{n(n-1)}}$$

### Std Error of Standard Deviation

This is a large sample estimate of the standard error of *s* for an infinite population.

### LCL of Standard Deviation

This is the lower value of a 100(1-α) interval estimate for the standard deviation based on the chi-squared distribution with n-1 degrees of freedom. This interval estimate assumes that the variable is normally distributed.

$$LCL = \sqrt{\frac{s^2(n-1)}{\chi^2_{\alpha/2,n-1}}}$$

### UCL of Standard Deviation

This is the upper value of a 100(1-α) interval estimate for the standard deviation based on the chi-squared distribution with n-1 degrees of freedom. This interval estimate assumes that the variable is normally distributed.

$$UCL = \sqrt{\frac{s^2(n-1)}{\chi^2_{1-\alpha/2,n-1}}}$$

### Unbiased Std Dev

This is an unbiased estimate of the standard deviation. If the data come from a normal distribution, the sample variance, $s^2$, is an unbiased estimate of the population variance. Unfortunately, the sample standard deviation, $s$, is a biased estimate of the population standard deviation. This bias is usually overlooked, but division of $s$ by a correction factor, $c_4$, will correct for this bias. This is frequently done in quality control applications. The formula for $c_4$ is:

$$c_4 = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

where

$$\Gamma(n) = \int_0^\infty t^{n-1}e^{-t}dt$$

### Std Error of Mean

This is an estimate of the standard error of the mean. This is an estimate of the precision of the sample mean. It, its standard error and confidence limits, are calculated by dividing the corresponding Standard Deviation value by the square root of $n$.

### Interquartile Range

This is the interquartile range (IQR). It is the difference between the third quartile and the first quartile (between the 75th percentile and the 25th percentile). This represents the range of the middle 50 percent of the distribution. It is a very robust (not affected by outliers) measure of dispersion. In fact, if the data are normally distributed, a robust estimate of the sample standard deviation is IQR/1.35. If a distribution is very concentrated around its mean, the IQR will be small. On the other hand, if the data are widely dispersed, the IQR will be much larger.

### Range

The difference between the largest and smallest values for a variable. If the data for a given variable is normally distributed, a quick estimate of the standard deviation can be made by dividing the range by six.

## Skewness and Kurtosis Section

**Skewness and Kurtosis Section of Height**

| Parameter | Skewness | Kurtosis | Fisher's g1 | Fisher's g2 | Coefficient of Variation | Coefficient of Dispersion |
|---|---|---|---|---|---|---|
| Value | 0.471155 | 2.140641 | 0.5102501 | -0.7479873 | 0.135928 | 0.1142857 |
| Std Error | 0.3343679 | 0.5338696 | | | 0.0148992 | |

### Skewness

This statistic measures the direction and degree of asymmetry. A value of zero indicates a symmetrical distribution. A positive value indicates skewness (longtailedness) to the right while a negative value indicates skewness to the left. Values between -3 and +3 indicate are typical values of samples from a normal distribution. For an alternative measure of skewness, see *Fisher's g1*, below.

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}}$$

### Std Error of Skewness

This is a large sample estimate of the standard error of skewness for an infinite population.

### Kurtosis

This statistic measures the heaviness of the tails of a distribution. The usual reference point in kurtosis is the normal distribution. If this kurtosis statistic equals three and the skewness is zero, the distribution is normal. Unimodal distributions that have kurtosis greater than three have heavier or thicker tails than the normal. These same distributions also tend to have higher peaks in the center of the distribution (leptokurtic). Unimodal distributions whose tails are lighter than the normal distribution tend to have a kurtosis that is less than three. In this case, the peak of the distribution tends to be broader than the normal (platykurtic). Be forewarned that this statistic is an unreliable estimator of kurtosis for small sample sizes. For an alternative measure of skewness, see *Fisher's g2*, below.

$$b_2 = \frac{m_4}{m_2^2}$$

### Std Error of Kurtosis

This is a large sample estimate of the standard error of skewness for an infinite population.

### Fisher's g1

Fisher's $g_1$ measure is an alternative measure of skewness.

$$g_1 = \frac{\sqrt{n(n-1)}\ b_1}{n-2}$$

### Fisher's g2

The Fisher's $g_2$ measure is an alternative measure of kurtosis.

$$g_2 = \frac{(n+1)(n-1)}{(n-2)(n-3)}\left[ b_2 - \frac{3(n-1)}{n+1} \right]$$

### Coefficient of Variation

The *coefficient of variation* is a relative measure of dispersion. It is most often used to compare the amount of variation in two samples. It can be used for the same data over two time periods or for the same time period but two different places. It is the standard deviation divided by the mean:

$$cv = \frac{s}{\bar{x}}$$

### Std Error of Coefficient of Variation

This is a large sample estimate of the standard error of the estimated coefficient of variation.

### Coefficient of Dispersion

The *coefficient of dispersion* is a robust, relative measure of dispersion. It is frequently used in real estate or tax assessment applications.

$$COD = \frac{\left(\frac{\sum |\, x_i - median \,|}{n}\right)}{median}$$

## Trimmed Section

**Trimmed Section of Height**

| Parameter | 5% Trimmed | 10% Trimmed | 15% Trimmed | 25% Trimmed | 35% Trimmed | 45% Trimmed |
|---|---|---|---|---|---|---|
| Trim-Mean | 61.77778 | 61.5 | 61.35714 | 60.9 | 60.5 | 59.5 |
| Trim-Std Dev | 7.448297 | 6.552353 | 5.692196 | 3.60401 | 2.428992 | 0.7071068 |
| Count | 18 | 16 | 14 | 10 | 6 | 2 |

### %Trimmed

We call *100g* the trimming percentage, the percent of data that is trimmed from each side of the sorted data. Thus, if $g$ = 5%, for a sample size of 200, 10 observations are ignored from each side of the sorted array of data values. Note that our formulation allows fractional data values. Different trimming percentages are available, but 5% and 10% are the most common in practice.

### Trim-Mean

These are the alpha-trimmed means discussed by Hoaglin (1983, page 311). These are useful for quickly assessing the impact of outliers. You would like to see stability in these trimmed means after a small degree of trimming. The formula for the trimmed mean for 100g% trimming is

$$\bar{x}_{(\alpha)} = \frac{1}{n(1-2\alpha)}\left\{(1-r)\left[X_{(g+1)} + X_{(n-g)}\right] + \sum_{i=g+2}^{n-g-1} X_{(i)}\right\}$$

where $g = \lfloor \alpha n \rfloor$ and $r = \alpha n - g$.

### Trim-Std Dev

This is the standard deviation of the observations that remain after the trimming. It can be used to evaluate changes in the standard deviation for different degrees of trimming. The formula for the trimmed standard deviation for *100g%* trimming is the standard formula for a weighted average using the weights given below.

$$a_i = 0 \text{ if } i \le g \text{ or } i \ge n - g + 1$$

$$a_i = \frac{1-r}{n-2\alpha n} \text{ if } i = g+1 \text{or } i = n-g$$

$$a_i = \frac{1}{n-2\alpha n} \text{ if } g+2 \le i \le n-g-1$$

### Count

This is the number of observations remaining after the trimming operation. Note that this may be a fractional amount under alpha-trimming.

# Mean-Deviation Section

**Mean-Deviation Section of Height**

| Parameter | \|X-Mean\| | \|X-Median\| | (X-Mean)^2 | (X-Mean)^3 | (X-Mean)^4 |
|---|---|---|---|---|---|
| Average | 7.01 | 6.8 | 67.69 | 262.392 | 9808.281 |
| Std Error | 1.134273 | | 16.16531 | 181.2807 | 3522.41 |

### Average of |X-Mean|

This is a measure of dispersion, called the *mean deviation* or the *mean absolute deviation*. It is not affected by outliers as much as the standard deviation, since the differences from the mean are not squared. If the distribution for the variable of interest is normal, the mean deviation is approximately equal to 0.8 standard deviations.

$$MAD = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

### Std Error of |X-Mean|

This is an estimate of the standard error of the *mean deviation*.

$$SE_{MAD} = \sqrt{\frac{2s^2(n-1)}{\pi n^2}\left[\frac{\pi}{2} + (n^2 - 2n)^2 - n + \arcsin\left(\frac{1}{n-1}\right)\right]}$$

### Average of |X-Median|

This is an alternate formulation of the *mean deviation* above that is more robust to outliers since the median is used as the center point of the distribution.

$$MAD_{\text{Robust}} = \frac{\sum_{i=1}^{n} |x_i - \text{median}|}{n}$$

### Average of (X-Mean)^2

This is the second moment about the mean, $m_2$.

### Std Error of (X-Mean)^2

This is the estimated standard deviation of the second moment.

### Average of (X-Mean)^3

This is the third moment about the mean, $m_3$.

### Std Error of (X-Mean)^3

This is the estimated standard deviation of the third moment.

### Average of (X-Mean)^4

This is the fourth moment about the mean, $m_4$.

### Std Error of (X-Mean)^4

This is the estimated standard deviation of the fourth moment.

## Quartile Section

This gives the value of the $j^{th}$ percentile. Of course, the $25^{th}$ percentile is called the *first (lower) quartile*, the $50^{th}$ percentile is the *median*, and the $75^{th}$ percentile is called the *third (upper) quartile*.

**Quartile Section**

| Parameter | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
|---|---|---|---|---|---|
| Value | 52 | 56 | 59.5 | 70 | 75.7 |
| 95% LCL | | 51 | 56 | 60 | |
| 95% UCL | | 59 | 67 | 76 | |

### Value

These are the values of the specified percentiles. Note that the definition of a percentile depends on the type of percentile that was specified.

### LCL and 95% UCL

These give an exact, 100(1-α)% confidence interval for the population percentile. This confidence interval does not assume normality. Instead, it only assumes a random sample of *n* items from a continuous distribution. The interval is based on the equation:

$$1-\alpha = I_p(r, n-r+1) - I_p(n-r+1, r)$$

Here $I_p(a,b)$ is the integral of the incomplete beta function:

$$I_q(n-r+1, r) = \sum_{k=0}^{r-1} \binom{n}{k} p^k (1-p)^{n-k}$$

and q=1-p and $I_p(a,b) = 1 - I_{1-p}(b,a)$.

## Normality Test Section

**Normality Test Section**

| Test Name | Test Value | Prob Level | 10% Critical Value | 5% Critical Value | Decision (5%) |
|---|---|---|---|---|---|
| Shapiro-Wilk W | 0.9373675 | 0.213730 | | | Can't reject normality |
| Anderson-Darling | 0.4433714 | 0.286286 | | | Can't reject normality |
| Martinez-Iglewicz | 1.025854 | | 1.216194 | 1.357297 | Can't reject normality |
| Kolmogorov-Smirnov | 0.1482353 | | 0.176 | 0.192 | Can't reject normality |
| D'Agostino Skewness | 1.0367 | .299858 | 1.645 | 1.960 | Can't reject normality |
| D'Agostino Kurtosis | -.7855 | .432156 | 1.645 | 1.960 | Can't reject normality |
| D'Agostino Omnibus | 1.6918 | .429161 | 4.605 | 5.991 | Can't reject normality |

### Normality Tests

This section displays the results of seven tests of the hypothesis that the data come from the normal distribution. The Shapiro-Wilk andAnderson-Darling tests are usually considered as the best. The Kolmogorov-Smirnov test is included because of its historical popularity, but is bettered in almost every way by the other tests.

Unfortunately, these tests have small statistical power (probability of detecting nonnormal data) unless the sample sizes are large, say over 100. Hence, if the decision is to reject, you can be reasonably certain that the data are not normal. However, if the decision is to accept, the situation is not as clear. If you have a sample size of 100 or more, you can reasonably assume that the

actual distribution is closely approximated by the normal distribution. If your sample size is less than 100, all you know is that there was not enough evidence in your data to reject the normality assumption. In other words, the data might be nonnormal, you just could not prove it. In this case, you must rely on the graphics and past experience to justify the normality assumption.

### Shapiro-Wilk W Test

This test for normality, developed by Shapiro and Wilk (1965), has been found to be the most powerful test in most situations. It is the ratio of two estimates of the variance of a normal distribution based on a random sample of n observations. The numerator is proportional to the square of the best linear estimator of the standard deviation. The denominator is the sum of squares of the observations about the sample mean. W may be written as the square of the Pearson correlation coefficient between the ordered observations and a set of weights which are used to calculate the numerator. Since these weights are asymptotically proportional to the corresponding expected normal order statistics, W is roughly a measure of the straightness of the normal quantile-quantile plot. Hence, the closer W is to one, the more normal the sample is.

The probability values for W are valid for samples in the range of 3 to 5000.

W may not be as powerful as other tests when ties occur in your data.

The test is not calculated when a frequency variable is specified.

### Anderson-Darling Test

This test, developed by Anderson and Darling (1954), is the most popular normality test that is based on EDF statistics. In some situations, it has been found to be as powerful as the Shapiro-Wilk test.

The test is not calculated when a frequency variable is specified.

### Martinez-Iglewicz

This test for normality, developed by Martinez and Iglewicz (1981), is based on the median and a robust estimator of dispersion. They have shown that this test is very powerful for heavy-tailed symmetric distributions as well as a variety of other situations. A value of the test statistic that is close to one indicates that the distribution is normal. This test is recommended for exploratory data analysis by Hoaglin (1983). The formula for this test is:

$$I = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{(n-1)s_{bi}^2}$$

where $s_{bi}^2$ is a biweight estimator of scale.

### Martinez-Iglewicz (10% Critical and 5% Critical)

The 10% and 5% critical values are given here. If the value of the test statistic is greater than this value, reject normality at that level of significance.

### Martinez-Iglewicz Decision (5%)

This reports the outcome of this test at the 5% significance level.

### Kolmogorov-Smirnov

This test for normality is based on the maximum difference between the observed distribution and expected cumulative-normal distribution. Since it uses the sample mean and standard deviation to calculate the expected normal distribution, the Lilliefors' adjustment is used. The smaller the maximum difference the more likely that the distribution is normal.

This test has been shown to be less powerful than the other tests in most situations. It is included because of its historical popularity.

### Kolmogorov-Smirnov (10% Critical and 5% Critical)

The 10% and 5% critical values are given here. If the value of the test statistic is greater than this value, reject normality at that level of significance. The critical values are the Lilliefors' adjusted values as given by Dallal (1986). If the test value is greater than the reject critical value, normality is rejected at that level of significance.

### Kolmogorov-Smirnov Decision (5%)

This reports the outcome of this test at the 5% significance level.

### D'Agostino Skewness

D'Agostino (1990) describes a normality test based on the skewness coefficient, $\sqrt{b_1}$. Recall that because the normal distribution is symmetrical, $\sqrt{b_1}$ is equal to zero for normal data. Hence, a test can be developed to determine if the value of $\sqrt{b_1}$ is significantly different from zero. If it is, the data are obviously nonnormal. The statistic, $z_s$, is, under the null hypothesis of normality, approximately normally distributed. The computation of this statistic, which is restricted to sample sizes $n>8$, is

$$z_s = d \ln\left(\frac{T}{a} + \sqrt{\left(\frac{T}{a}\right)^2 + 1}\right)$$

where

$$b_1 = \frac{m_3^2}{m_2^3}$$

$$T = \sqrt{b_1\left(\frac{(n+1)(n+3)}{6(n-2)}\right)}$$

$$C = \frac{3(n^2 + 27n - 70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}$$

$$W^2 = -1 + \sqrt{2(C-1)}$$

$$a = \sqrt{\frac{2}{W^2 - 1}}$$

$$d = \frac{1}{\sqrt{\ln(W)}}$$

### Skewness Test (Prob Level)

This is the two-tail, significance level for this test. Reject the null hypothesis of normality if this value is less than a pre-determined value, say 0.05.

### Skewness Test Decision (5%)

This reports the outcome of this test at the 5% significance level.

### D'Agostino Kurtosis

D'Agostino (1990) describes a normality test based on the kurtosis coefficient, $b_2$. Recall that for the normal distribution, the theoretical value of $b_2$ is 3. Hence, a test can be developed to determine if the value of $b_2$ is significantly different from 3. If it is, the data are obviously nonnormal. The statistic, $z_k$, is, under the null hypothesis of normality, approximately normally distributed for sample sizes n>20. The calculation of this test proceeds as follows:

$$z_k = \frac{\left(1-\frac{2}{9A}\right) - \left(\frac{1-\frac{2}{A}}{1+G\sqrt{\frac{2}{A-4}}}\right)^{1/3}}{\sqrt{\frac{2}{9A}}}$$

where

$$b_2 = \frac{m_4}{m_2^2}$$

$$G = \frac{b_2 - \left(\frac{3n-3}{n+1}\right)}{\sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}}$$

$$E = \frac{6(n^2-5n+2)}{(n+7)(n+9)}\sqrt{\frac{6(n+3)(n+5)}{n(n-2)(n-3)}}$$

$$A = 6 + \frac{8}{E}\left(\frac{2}{E} + \sqrt{1+\frac{4}{E^2}}\right)$$

### Prob Level of Kurtosis Test

This is the two-tail significance level for this test. Reject the null hypothesis of normality if this value is less than a pre-determined value, say 0.05.

### Decision of Kurtosis Test

This reports the outcome of this test at the 5% significance level.

### D'Agostino Omnibus

D'Agostino (1990) describes a normality test that combines the tests for skewness and kurtosis. The statistic, $K^2$, is approximately distributed as a chi-square with two degrees of freedom. After calculated $z_s$ and $z_k$, calculate $K^2$ as follows:

$$K^2 = z_s^2 + z_k^2$$

### Prob Level D'Agostino Omnibus

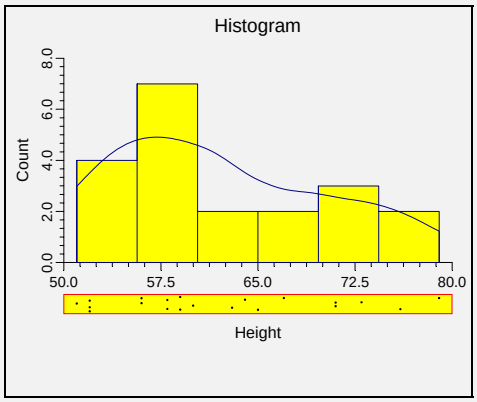
This is the significance level for this test. Reject the null hypothesis of normality if this value is less than a pre-determined value, say 0.05.

### Decision of D'Agostino Omnibus Test

This reports the outcome of this test at the 5% significance level.

## Histogram Plot

The following plot combines a histogram, a density trace, and a dot plot.




### Histogram

The histogram is a traditional way of displaying the shape of a batch of data. It is constructed from a frequency distribution, where choices on the number of classes and class width have been made. These choices can drastically affect the shape of the histogram. The ideal shape to look for in the case of normality is a bell-shaped symmetrical distribution.

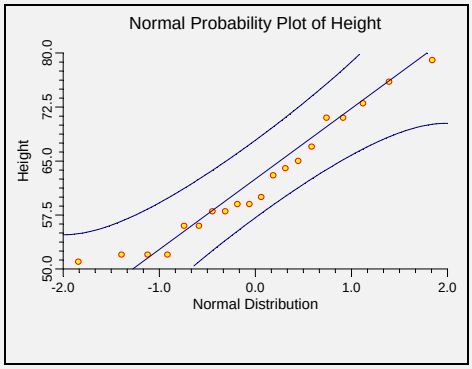
### Density Trace

The density trace is a smoothed histogram in which the class width or interval and the number of bins or classes does not bias the perspective of shape. It is generally overlaid on top of the histogram. In evaluating normality, we look for a bell-shaped symmetrical distribution.

### Dot Plot

This plot displays the data along the horizontal axis. A random, vertical component is added so that two points are not plotted at exactly the same point. The dot plot reminds you of the pattern of the actual data going into the histogram and density trace.

## Normal Probability Plot




This is a plot of the inverse of the standard normal cumulative versus the ordered observations. If the underlying distribution of the data is normal, the points will fall along a straight line. Deviations from this line correspond to various types of nonnormality. Stragglers at either end of the normal probability plot indicate outliers. Curvature at both ends of the plot indicates long or short distribution tails. Convex, or concave, curvature indicates a lack of symmetry. Gaps, plateaus, or segmentation in the plot indicate certain phenomenon that need closer scrutiny.

Confidence bands serve as a visual reference for departures from normality. If any of the observations fall outside the confidence bands, the data are not normal. The numerical normality tests will usually confirm this fact statistically. If only one observation falls outside the confidence limits, it may be an outlier. Note that these confidence bands are based on large sample formulas. They may not be accurate for small samples (less than 30).

## Percentile Section

**Percentile Section of Height**

| Percentile | Value | 95% LCL | 95% UCL | Exact Conf. Level |
|---|---|---|---|---|
| 99 | 79 | | | |
| 95 | 78.85 | | | |
| 90 | 75.7 | | | |
| 85 | 72.7 | 64 | 79 | 95.53193 |
| 80 | 71 | 64 | 79 | 95.63281 |
| 75 | 70 | 60 | 76 | 96.1823 |
| 70 | 66.4 | 59 | 76 | 97.52179 |
| 65 | 64.65 | 59 | 73 | 96.83029 |
| 60 | 63.6 | 58 | 71 | 96.30099 |
| 55 | 61.65 | 58 | 71 | 95.97224 |
| 50 | 59.5 | 56 | 67 | 95.86105 |
| 45 | 59 | 56 | 65 | 95.97224 |
| 40 | 58.4 | 52 | 64 | 96.30099 |
| 35 | 58 | 52 | 63 | 96.83029 |
| 30 | 56.6 | 52 | 60 | 97.52179 |
| 25 | 56 | 51 | 59 | 95.59036 |
| 20 | 52.8 | 51 | 58 | 95.63281 |
| 15 | 52 | 51 | 58 | 95.53193 |
| 10 | 52 | | | |
| 5 | 51.05 | | | |
| 1 | 51 | | | |

Percentile Formula: Ave X(p[n+1])

This section gives a larger set of percentiles than was included in the Quartile Section. Use it when you need a less common percentile.

### Percentile

This is the percentage amount that you want the percentile of.

### Value

This gives the value of the $p^{th}$ percentile. Note that the percentile method used is listed at the bottom of the report.

### 95%LCL and 95% UCL

These give an exact, 100(1-α)% confidence interval for the population percentile. This confidence interval does not assume normality. Instead, it only assumes a random sample of *n* items from a continuous distribution. The interval is based on the equation:

$$1-\alpha = I_p(r, n-r+1) - I_p(n-r+1, r)$$

Here $I_p(a,b)$ is the integral of the incomplete beta function:

$$I_q(n-r+1, r) = \sum_{k=0}^{r-1} \binom{n}{k} p^k (1-p)^{n-k}$$

and q=1-p and $I_p(a,b) = 1- I_{1-p}(b,a)$.

### Exact Conf. Level

Because of the discrete nature of the confidence interval constructed above, ***NCSS*** finds an interval that is less than the specified alpha level. This column gives the actual confidence coefficient of the interval.

## Stem-Leaf Plot Section

**Stem-Leaf Plot Section of Height**

| Depth | Stem | | Leaves |
|---|---|---|---|
| 4 | 5* | \| | 1222 |
| 10 | . | \| | 668899 |
| 10 | 6* | \| | 034 |
| 7 | . | \| | 57 |
| 5 | 7* | \| | 113 |
| 2 | . | \| | 69 |

Unit = 1 Example: 1|2 Represents 12

The stem-leaf plot is a type of histogram which retains much of the identity of the original data. It is useful for finding data-entry errors as well as for studying the distribution of a variable.

### Depth

This is the cumulative number of leaves, counting in from the nearest end.

### Stem

The stem is the first digit of the actual number. For example, the stem of the number 523 is 5 and the stem of 0.0325 is 3. This is modified appropriately if the batch contains numbers of different orders of magnitude. The largest order of magnitude is used in determining the stem. Depending upon the number of leaves, a stem may be divided into two or more sub-stems. A special set of symbols is then used to mark the stems.

In the current example, the star (*) represents numbers in the range of zero to four, while the period (.) represents numbers in the range of five to nine.

### Leaf

The leaf is the second digit of the actual number. For example, the leaf of the number 523 is 2 and the leaf of 0.0325 is 2. This is modified appropriately if the batch contains numbers of different orders of magnitude. The largest order of magnitude is used in determining the leaf.

### Unit

This line at the bottom indicates how the data were scaled to make the plot.

## Chapter 201

# Descriptive Tables

## Introduction

This procedure produces tables of means, medians, standard deviations, coefficients of variation, sums, and counts for various combinations of control (break) variables. Seven tabular formats are available. The tables are similar in structure to those produced by cross tabulation.

This module is used to summarize data containing a combination of continuous and categorical variables. Large volumes of such data may be summarized in tables of means, counts, or standard deviation. Discussions of these statistics may be found in the Descriptive Statistics chapter and will not be reproduced here.

## Types of Categorical Variables

Note that we will refer to two types of categorical variables: *By* and *Break***.** Break variables are used to split a database into subgroups. A separate table is generated for each unique set of values of the Break variables. The values of a By variable are used to define the rows and columns of the tabulation table. Up to two By variables may be used per table.

## Data Structure

The data below are a subset of the RESALE database provided with the software. This (computer simulated) data gives the selling price, the number of bedrooms, the total square footage (finished and unfinished), and the size of the lots for 150 residential properties sold during the last four months in two states. Only the first 8 of the 150 observations are displayed.

**RESALE dataset (subset)**

| State | Price | Bedrooms | TotalSqft | LotSize |
|---|---|---|---|---|
| Nev | 260000 | 2 | 2042 | 10173 |
| Nev | 66900 | 3 | 1392 | 13069 |
| Vir | 127900 | 2 | 1792 | 7065 |
| Nev | 181900 | 3 | 2645 | 8484 |
| Nev | 262100 | 2 | 2613 | 8355 |
| Nev | 147500 | 2 | 1935 | 7056 |
| Nev | 167200 | 2 | 1278 | 6116 |
| Nev | 395700 | 2 | 1455 | 14422 |

# Missing Values

The treatment of missing values must be carefully considered. You have the option to ignore missing values completely or to include them in the reports. If they are ignored, observations with missing values in either the categorical variable or the continuous variable are removed.

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

## Variables Tab

This panel specifies the variables that will be used in the analysis.

You can specify a *Table Column* variable or a *Table Row* variable or both. The unique values of these two variables will form the columns and rows of the table. If more than one variable is specified in either section, a separate table will be generated for each combination of variables.

Four types of categorical variables may be specified:

1. Variables containing text values. These are called *Discrete Variables*.
2. Variables containing numeric values that are to be treated individually. For example, you might have used a set of index numbers like “1 2 3 4” to represent four states. These are also called *Discrete Variables.*
3. Variables containing numeric values that are to be grouped or combined into a set of predefined intervals. *You specify the interval boundaries*. For example, a variable containing age values might be grouped as “Under 21, 21 to 55, and Over 55.” The key is that you specify the intervals. These are called *Numeric Variables (Limits).*
4. Variables containing numeric values that are to be combined into a set of computer-generated intervals. *You specify only the number of intervals*. The program determines a set of equal-length intervals based on the minimum and maximum found in the data. This format may cause problems since you do not set the interval boundaries directly. These are called *Numeric Variables (Width).*

### Data Variables

#### Response Variables

Select at least one response variable. The statistics (means, standard deviations, etc.) generated will be for the values in these variables.

### Frequency Variable

#### Frequency Variable

This optional variable specifies the number of observations that each row represents. When omitted, each row represents a single observation. If your data is the result of previous summarization, you may want certain rows to represent several observations. Note that negative

values are treated as a zero frequency and are omitted. Fractional values may be used. You may also think of this as a weighting variable.

## Select Table Type

### Table Format

This option specifies which of the seven table formats you want to use. These formats were created based on the number of By variables used (0, 1, or 2), the number of Response Variables displayed, and the number of statistics displayed.

- **1 Combined Stats, No By's**

  A single row of the specified statistics (count, mean, etc.) is generated for each Response Variable. Any Table Column Variables or Table Row Variables specified are ignored. An example of this table format is:

**Table of Summary Statistics**

| Variable | Count | Mean | Median | Std Deviation |
|---|---|---|---|---|
| X1 | xxx | xxx | xxx | xxx |
| X2 | xxx | xxx | xxx | xxx |
| X3 | xxx | xxx | xxx | xxx |
| X4 | xxx | xxx | xxx | xxx |

- **2 Combined Stats, One By**

  A cross-tabulation table is constructed in which one side (row or column) is made up of the Response Variables and the other is made up of the categories of the Table Column Variable (or Table Row Variable). Selected statistics are shown as individual rows of the table. One table is generated for each Table Column Variable and Table Row Variable. An example of this table format is:

**Table of Counts, Means, and Standard Deviations**

| | By Var 1 | | | |
|---|---|---|---|---|
| Variables | Bv1 | Bv2 | Bv3 | Total |
| X1 | n | n | n | n |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| X2 | xxx | xxx | xxx | xxx |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| X3 | xxx | xxx | xxx | xxx |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| X4 | xxx | xxx | xxx | xxx |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |

- **3 Separate Stats, One By (Plots Possible)**

  A cross-tabulation table is constructed in which one side (row or column) is made up of the Response Variables and the other is made up of the categories of the Table Column Variable (or Table Row Variable). A separate table is generated for each statistic. One table is generated for each Table Column Variable and Table Row Variable. Examples of this table format are:

  **Table of Means**

| | By Var 1 | | | |
|---|---|---|---|---|
| Variables | Bv1 | Bv2 | Bv3 | Total |
| X1 | mean | mean | mean | mean |
| X2 | mean | mean | mean | mean |
| X3 | mean | mean | mean | mean |
| X4 | mean | mean | mean | mean |

  **Table of Std Deviations**

| | By Var 1 | | | |
|---|---|---|---|---|
| Variables | Bv1 | Bv2 | Bv3 | Total |
| X1 | std dev | std dev | std dev | std dev |
| X2 | std dev | std dev | std dev | std dev |
| X3 | std dev | std dev | std dev | std dev |
| X4 | std dev | std dev | std dev | std dev |

- **4 Combined Y's, Two By's**

  A cross-tabulation table is constructed in which columns are based on the Table Column Variable, the rows are based on the Table Row Variable, and a separate table row is given for each Response Variable. A single table is generated for each statistic (mean, count, etc.). An example of this table format is:

  **Table of Means**

| | By Var 1 | | | |
|---|---|---|---|---|
| By Var 2 | Bc1 | Bc2 | Bc3 | Total |
| Br1 | mean of X1 | mean of X1 | mean of X1 | mean of X1 |
| | mean of X2 | mean of X2 | mean of X2 | mean of X2 |
| | mean of X3 | mean of X3 | mean of X3 | mean of X3 |
| Br2 | mean of X1 | mean of X1 | mean of X1 | mean of X1 |
| | mean of X2 | mean of X2 | mean of X2 | mean of X2 |
| | mean of X3 | mean of X3 | mean of X3 | mean of X3 |
| Br3 | mean of X1 | mean of X1 | mean of X1 | mean of X1 |
| | mean of X2 | mean of X2 | mean of X2 | mean of X2 |
| | mean of X3 | mean of X3 | mean of X3 | mean of X3 |
| Total | mean of X1 | mean of X1 | mean of X1 | mean of X1 |
| | mean of X2 | mean of X2 | mean of X2 | mean of X2 |
| | mean of X3 | mean of X3 | mean of X3 | mean of X3 |

- **5 Combined Stats, Two By's**

  A cross-tabulation table is constructed in which columns are based on the Table Column Variable, the rows are based on the Table Row Variable, and a separate table row is given for each statistic selected (mean, count, etc.). A single table is generated for each Response Variable. An example of this table format is:

**Table of Counts, Means, and Standard Deviations of X1**

| | By Var 1 | | | |
|---|---|---|---|---|
| By Var2 | Bc1 | Bc2 | Bc3 | Total |
| Br1 | n | n | n | n |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| Br2 | n | n | n | n |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| Br3 | n | n | n | n |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |
| Total | n | n | n | n |
| | mean | mean | mean | mean |
| | std dev | std dev | std dev | std dev |

- **6 Separate Stats, Two By's (Plots Possible)**

  A cross-tabulation table is constructed in which columns are based on the Table Column Variable, the rows are based on the Table Row Variable, and a separate table is given for each combination of statistic (mean, count, etc.) and Response Variable. An example of this table format is:

**Table of Means of X1**

| | By Var 1 | | | |
|---|---|---|---|---|
| By Var2 | Bc1 | Bc2 | Bc3 | Total |
| Br1 | mean | mean | mean | mean |
| Br2 | mean | mean | mean | mean |
| Br3 | mean | mean | mean | mean |
| Total | mean | mean | mean | mean |

- **7 List Format, One Row-By**

  This format type creates a simple list of the data. This format requires that one Table-Row Variable be specified. You can also specify one or more Break Variables. The statistics listed on the report are specified by checking the appropriate checkboxes on this panel.

  This format is especially useful for creating a summarized version of a database. Here's how:

  1. Run this procedure selecting this type of table format.
  2. Copy the output to the Windows clipboard.
  3. Paste the information into a new datasheet. You will want to adjusted the variable names appropriately.

An example of this table format is:

**Summary List**

| Break1 | Break2 | ByVar1 | Count | Mean | StdDev |
|---|---|---|---|---|---|
| Bk1.1 | Bk2.1 | Bv1 | count | mean | stddev |
| Bk1.1 | Bk2.1 | Bv2 | count | mean | stddev |
| Bk1.1 | Bk2.1 | Bv3 | count | mean | stddev |
| Bk1.1 | Bk2.2 | Bv1 | count | mean | stddev |
| Bk1.1 | Bk2.2 | Bv2 | count | mean | stddev |
| Bk1.1 | Bk2.2 | Bv3 | count | mean | stddev |
| Bk1.2 | Bk2.1 | Bv1 | count | mean | stddev |
| Bk1.2 | Bk2.1 | Bv2 | count | mean | stddev |
| Bk1.2 | Bk2.1 | Bv3 | count | mean | stddev |

## 'By' Variables for Use in Table Columns and Rows

### Discrete Variables

This option specifies those variables that contain text and numeric values that are to be treated as discrete variables (Types 1 or 2). Variables containing text values are always listed here. Variables containing numeric values are listed here if you want each unique value to be treated separately.

### Numeric Variables (Width)

Use this option to specify variables that contain numeric values that are to be combined into a set of computer-generated intervals (Type 4). The intervals are specified in the three boxes: Number, Minimum, and Width. Note that you can specify one, two, or all three of these options.

### Number

The number of intervals to be created. If not enough intervals are specified to reach the maximum data value, more intervals are added.

### Minimum

The minimum value or the left boundary of the first interval. This value must be less than the minimum data value.

### Width

This is the width of an interval. A data value X is in this interval if Lower Limit < X <= Upper Limit. If this is left blank, it is calculated from the Number, Minimum, and maximum data value.

### Numeric Variables (Limits)

This specifies those variables that contain numeric values that are to be combined into a set of user-specified intervals (Type 3). The interval boundaries are specified as a list in the Interval Upper Limits box.

### Interval Upper Limits

Specify the upper limits of the intervals, separated by commas. For example, you would enter “1 3 5” to specify the four intervals: Under 1, 1 to 3, 3 to 5, and Over 5.

The logic structure of the interval is:

**Lower Bound < Value <= Upper Bound.**

Note that a “1” would be included in the “Under 1” interval, not the “1 to 3” interval. Also, a “5” would be included in the “3 to 5” interval, not the “Over 5” interval.

# Breaks Tab

This panel lets you specify up to eight break variables.

## Select Break (Grouping) Variables

### Break Variables

Specify one or more categorical variables whose distinct values will cause separate reports to be generated. Note that a separate set of reports (tables and plots) is generated for each unique set of values of these variables. Do not confuse these variables with the *Table Column* and *Table Row* variables, which specify the variables whose values will appear along the rows or columns of a particular table.

# Missing Tab

This panel lets you specify up to five missing values (besides the default of blank). For example, ‘0’, ‘9’, or ‘NA’ may be missing values in your database.

## Missing Value Options

### Missing Values

Specify up to five missing values here.

### Missing Value Inclusion

Specifies whether to include observations with missing values in the tables.

*Delete All* indicates that you want the missing values totally ignored.

*Include in Counts* indicates that you want the number of missing values displayed, but you do not want them to influence any of the percentages.

*Include in All* indicates that you want the missing values treated just like any other category. They will be included in all percentages and counts.

# Format Tab

The following options control the format of the reports.

## Format Options

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Value Labels

This option lets you select whether to display only values, value labels, or both. Use this option if you want the table to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Show Total

Specify whether to show row and/or column total statistics for those reports that use a by (Table Row or Table Column) variable.

### Label Justification

This option specifies whether the labels should be right or left justified above each column.

### Data Justification

This option specifies whether the data should be right or left justified in each cell.

### Split Column Headings

This option lets you select whether to split the column headings into two headings instead of one.

### Double Space

This option lets you select whether to insert an extra line at the end of each row section.

## Tabs

These options let you specify the tab settings across the table. The output ruler is also modified by the settings of *Label Justification* and *Data Justification*.

### First

Specifies the position of the first cell in inches. Note that the left-hand label always begins at 0.5 inches. Hence, the distance between this tab and 0.5 is the width provided for the row label information.

### Maximum

Specifies the right border of the table. The number of tabs is determined based on *First*, the *Increment*, and this option. If you set this value too large, your table may not be printed correctly.

### Increment

Specifies the width of a cell in inches.

### Offset

The amount (inches) of offset to the right used with a decimal tab on a custom ruler so the data is aligned properly under the left-justified column labels.

### Decimal Places

These options let you specify the number of decimal places used in the various items of the table.

#### Column-By

Specifies the number of decimal places displayed in the numeric *Table Columns* variable values. Note that *All* displays a single-precision (seven place accuracy).

#### Row-By

Specifies the number of decimal places displayed in the numeric *Table Rows* variable values. Note that *All* displays a single-precision (seven place accuracy).

#### Counts ... Maximums

Specifies the number of decimal places displayed in each statistic. Note that *All* displays the default amount.

## Reports Tab

These options control which of the available statistics are displayed on reports and plots.

### Select Statistics to be Displayed in Reports and Plots

#### Counts ... Standard Errors

For each of these statistics, specify whether you want a numeric report, a plot, or both.

## Plot Options Tab

The options on this panel control the appearance of the scatter plots of the statistics that may be displayed.

### Vertical and Horizontal Axis

#### Label

This is the text of the axis labels. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

#### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Line(s)

Specifies whether connect the points with lines for easier interpretation of trends.

## Plot Settings – Legend

### Show Legend

Specifies whether to display the legend.

### Legend Text

Specifies legend label. A {G} is replaced by the appropriate default value.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Show Break as Title

Specifies whether the current values of any *Break* variables should be displayed as a second title line in the plot.

# Symbols Tab

Specify the symbols used for each of the groups on the plots.

## Plotting Symbols

### Group 1-15

Specify the symbol used to designate a particular group. Double-click on a symbol or click on the button to the right of a symbol to specify the symbol’s size, type, and color.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Combined Stats, No By's

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price** to **LotSize** from the list of variables and then click **Ok**. "Price-LotSize" will appear in the Response Variables box.

**4 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Report**.
- In Means, select **Report**.

- In Medians, select **Report**.
- In Standard Deviations, select **Report**.
- In Sums, select **Report**.
- In COVs, select **Report**.
- In CODs, select **Report**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following report will be displayed in the Output window.

## Combined Stats, No By's Report

**Variable Summary Section**

| Variables | Count | Mean | Median | Standard Deviation | Sum |
|---|---|---|---|---|---|
| Price | 150 | 174392 | 158200 | 97656.81 | 2.61588E+07 |
| Year | 150 | 1971.273 | 1973 | 13.84667 | 295691 |
| Bedrooms | 150 | 2.42 | 2 | .8919476 | 363 |
| Bathrooms | 150 | 2.4 | 2.5 | .8047677 | 360 |
| Garage | 150 | 1.266667 | 1 | .5636252 | 190 |
| TotalSqft | 150 | 1893.38 | 1872.5 | 754.2496 | 284007 |
| LotSize | 150 | 8366.913 | 8344.5 | 2376.334 | 1255037 |

| Variables | COV | COD |
|---|---|---|
| Price | 0.55998 | 49.050 |
| Year | 0.00702 | 0.572 |
| Bedrooms | 0.36857 | 35.000 |
| Bathrooms | 0.33532 | 24.800 |
| Garage | 0.44497 | 36.000 |
| TotalSqft | 0.39836 | 31.980 |
| LotSize | 0.28402 | 23.993 |

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

## Example 2 – Combined Stats, One By

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price, TotalSqft,** and **LotSize** from the list of variables and then click **Ok**. "Price,TotalSqft,LotSize" will appear in the Response Variables box.
- In Table Format, select **2 Combined Stats, One By**.
- Double-click in the **'By' Variables for Use in Table Columns - Discrete Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. "State" will appear in the 'By' Variables for Use in Table Columns - Discrete Variables box.

**4 Specify the report format.**

- Click on the **Format tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- Check the box next to **Double Space**.
- In Tabs - First, enter **2.0**.

**5 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Report**.
- In Means, select **Report**.
- In Standard Deviations, select **Report**.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Combined Stats, One By Report

**Table of Counts, Means, Standard Deviations**

| | State | | |
|---|---|---|---|
| **Variables** | **Nevada** | **Virginia** | **Total** |
| Sales Price | 88 | 62 | 150 |
| | 170762.5 | 179543.5 | 174392 |
| | 98665.72 | 96771.49 | 97656.81 |
| Total Area (Sqft) | 88 | 62 | 150 |
| | 1881.33 | 1910.484 | 1893.38 |
| | 788.569 | 708.6572 | 754.2496 |
| Lot Size (Sqft) | 88 | 62 | 150 |
| | 8571.454 | 8076.597 | 8366.913 |
| | 2419.88 | 2301.226 | 2376.334 |

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

# Example 3 – Separate Stats, One By

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example3** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Bedrooms, Bathrooms, Garage,** and **Fireplace** from the list of variables and then click **Ok**. “Bedrooms-Fireplace” will appear in the Response Variables box.
- In Table Format, select **3 Separate Stats, One By**.
- Double-click in the **'By' Variables for Use in Table Columns - Discrete Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the 'By' Variables for Use in Table Columns - Discrete Variables box.

**4 Specify the report format.**

- Click on the **Format tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- In Show Total, select **On Reports and Plots**.

**5 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Omit**.
- In Means, select **Both**.
- In Standard Deviations, select **Omit**.

**6 Run the procedure.**

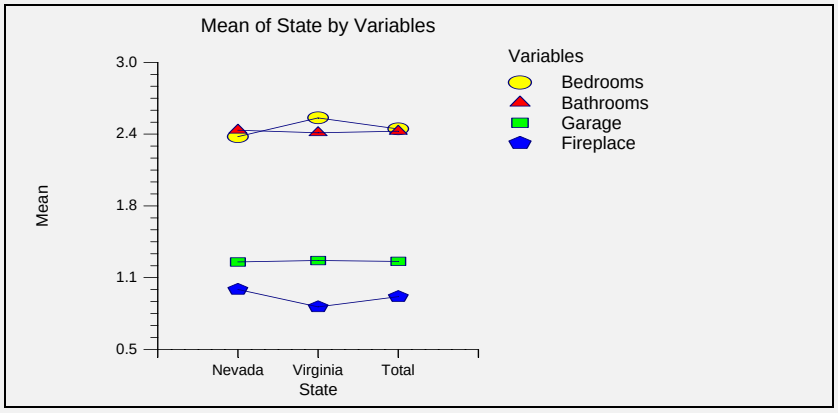
- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following report will be displayed in the Output window.

## Separate Stats, One By Report and Plot

**Table of Means**

| | State | | |
|---|---|---|---|
| **Variables** | **Nevada** | **Virginia** | **Total** |
| Bedrooms | 2.352273 | 2.516129 | 2.42 |
| Bathrooms | 2.409091 | 2.387097 | 2.4 |
| Garage | 1.261364 | 1.274194 | 1.266667 |
| Fireplace | 1.022727 | .8709677 | .96 |




The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

# Example 4 – Combined Y's, Two By's

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example4** from the Template tab of the Descriptive Tables window.

### 1 Open the RESALE dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

### 2 Open the Descriptive Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price, FinishSqft,** and **LotSize** from the list of variables and then click **Ok**. "Price,FinishSqft-LotSize" will appear in the Response Variables box.

- In Table Format, select **4 Combined Y's, Two Bys**.
- Double-click in the **'By' Variables for Use in Table Columns - Discrete Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. "State" will appear in the Table Columns - Discrete Variables box.
- Double-click in the **'By' Variables for Use in Table Rows - Numeric Variables (Limits)** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. "TotalSqft" will appear in the 'By' Variables for Use in Table Rows - Numeric Variables (Limits) box.
- In 'By' Variables for Use in Table Rows - Interval Upper Limits, enter **1000 2000 3000**.

**4 Specify the report format.**

- Click on the **Format tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- In Show Total, select **On Reports and Plots**.
- Check Double Space.
- In Tabs - First, enter **2.0**.

**5 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Omit**.
- In Means, select **Report**.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following report will be displayed in the Output window.

## Combined Y's, Two By's Report

**Means of Sales Price, Finished Area (Sqft), Lot Size (Sqft)**

| | State | | |
|---|---|---|---|
| **Total Area (Sqft)** | **Nevada** | **Virginia** | **Total** |
| Under 1000 | 160475 | 142850 | 152921.4 |
| | 738.125 | 739.6667 | 738.7857 |
| | 8816 | 9857.833 | 9262.5 |
| 1000 To 2000 | 153293.3 | 172992.9 | 160849.3 |
| | 1234.311 | 1247.179 | 1239.247 |
| | 9094.8 | 7674.286 | 8549.945 |
| 2000 To 3000 | 197200 | 186461.5 | 192029.6 |
| | 1974.214 | 2086.077 | 2028.074 |
| | 7503.179 | 8129.808 | 7804.889 |
| Over 3000 | 189071.4 | 291400 | 211811.1 |
| | 3375.143 | 2871 | 3263.111 |
| | 9200.714 | 7673.5 | 8861.333 |

| | | | |
|---|---|---|---|
| Total | 170762.5 | 179543.5 | 174392 |
| | 1594.92 | 1602.242 | 1597.947 |
| | 8571.454 | 8076.597 | 8366.913 |

The definitions of these statistics are identical to those found in the *Descriptive Statistics* chapter. They will not be repeated here.

## Example 5 – Combined Stats, Two By's

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example5** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. "Price" will appear in the Response Variables box.
- In Table Format, select **5 Combined Stats, Two Bys**.
- Double-click in the **'By' Variables for Use in Table Columns - Discrete Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. "State" will appear in the 'By' Variables for Use in Table Columns - Discrete Variables box.
- Double-click in the **'By' Variables for Use in Table Rows - Numeric Variables (Limits)** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. "TotalSqft" will appear in the 'By' Variables for Use in Table Rows - Numeric Variables (Limits) box.
- In 'By' Variables for Use in Table Rows - Interval Upper Limits, enter **1000 2000 3000**.

**4 Specify the report format.**

- Click on the **Format tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- In Show Total, select **On Reports Only**.
- Check Double Space.
- In Tabs - First, enter **2.0**.

**5 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Report**.
- In Means, select **Report**.
- In Medians, select **Report**.
- In Standard Deviations, select **Report**.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following report will be displayed in the Output window.

## Combined Stats, Two By's Report

**Table of Counts, Means, Medians, Standard Deviations of Sales Price**

| | State | | |
|---|---|---|---|
| **Total Area (Sqft)** | **Nevada** | **Virginia** | **Total** |
| Under 1000 | 8 | 6 | 14 |
| | 160475 | 142850 | 152921.4 |
| | 136050 | 85200 | 110500 |
| | 110945.7 | 107838.2 | 105747.5 |
| 1000 To 2000 | 45 | 28 | 73 |
| | 153293.3 | 172992.9 | 160849.3 |
| | 123400 | 163000 | 150100 |
| | 91336.91 | 71798.73 | 84405.74 |
| 2000 To 3000 | 28 | 26 | 54 |
| | 197200 | 186461.5 | 192029.6 |
| | 182850 | 145550 | 176250 |
| | 106136.7 | 111024.2 | 107621.7 |
| Over 3000 | 7 | 2 | 9 |
| | 189071.4 | 291400 | 211811.1 |
| | 150900 | 291400 | 168500 |
| | 94037.06 | 173806.8 | 111554.4 |
| Total | 88 | 62 | 150 |
| | 170762.5 | 179543.5 | 174392 |
| | 151050 | 162800 | 158200 |
| | 98665.72 | 96771.49 | 97656.81 |

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

# Example 6 – Separate Stats, Two By's

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example6** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. "Price" will appear in the Response Variables box.
- In Table Format, select **6 Separate Stats, Two Bys**.
- Double-click in the **'By' Variables for Use in Table Rows - Discrete Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. "State" will appear in the 'By' Variables for Use in Table Rows - Discrete Variables box.
- Double-click in the **'By' Variables for Use in Table Columns - Numeric Variables (Limits)** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. "TotalSqft" will appear in the 'By' Variables for Use in Table Columns - Numeric Variables (Limits) box.
- In 'By' Variables for Use in Table Columns - Interval Upper Limits, enter **1000 2000 3000**.

**4 Specify the report format.**

- Click on the **Format tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- In Show Total, select **On Reports and Plots**.
- In Label Justification, select **Right**.
- In Data Justification, select **Right**.
- Check Double Space.
- In Tabs - First, enter **2.0**.
- In Decimal Places - Means, enter **0**.

**5 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Omit**.
- In Means, select **Both**.

**6 Specify the plots.**

- Click on the Plot Options tab.
- Click on the Vertical Axis - Tick Label Settings button.
- In Decimals, select **0**.
- Click on **Ok** to close the settings window.
- Click on the Horizontal Axis - Tick Label Settings button.
- In Decimals, select **0**.
- Under Text Rotation, select **Vertical**.
- In Max Characters, select **15**.
- Click on **Ok** to close the settings window.

**7 Run the procedure.**

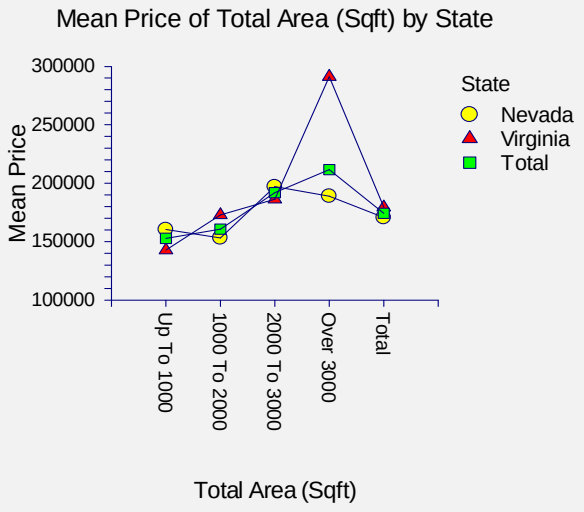
- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Separate Stats, Two By's Report and Plot

**Means of Sales Price**

**Total Area (Sqft)**

| State | Up To 1000 | 1000 To 2000 | 2000 To 3000 | Over 3000 | Total |
|---|---|---|---|---|---|
| Nevada | 160475 | 153293 | 197200 | 189071 | 170763 |
| Virginia | 142850 | 172993 | 186462 | 291400 | 179544 |
| Total | 152921 | 160849 | 192030 | 211811 | 174392 |



# Example 7 – List Format

The data used are found in the RESALE database. You may follow along here by making the appropriate entries or load the completed template **Example7** from the Template tab of the Descriptive Tables window.

**1 Open the RESALE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Resale.s0**.
- Click **Open**.

**2 Open the Descriptive Tables window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Tables**. The Descriptive Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Descriptive Tables window, select the **Variables tab**.
- Double-click in the **Response Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. "Price" will appear in the Response Variables box.
- In Table Format, select **7 List Format, One Row-By**.
- Double-click in the **'By' Variables for Use in Table Rows - Discrete Variables** text box. This will bring up the variable selection window.
- Select **Neighborhood** from the list of variables and then click **Ok**. "**Neighborhood**" will appear in the 'By' Variables for Use in Table Rows - Discrete Variables box.

**4 Specify the break variables.**

- On the Descriptive Tables window, select the **Breaks tab**.
- Double-click in the **first Break Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. "State" will appear in the first Break Variables box.
- Double-click in the **second Break Variables** text box. This will bring up the variable selection window.
- Select **City** from the list of variables and then click **Ok**. "City" will appear in the second Break Variables box.

**5 Specify the report format.**

- Click on the **Format tab**.
- In Show Total, select **Omit Totals**.

**6 Specify the reports.**

- Click on the **Reports tab**.
- In Counts, select **Report**.
- In Means, select **Report**.
- In Standard Deviations, select **Report**.

**7 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## List Format Report

**Summary List**

| State | City | Neighborhood | Price Count | Price Mean | Price StdDev |
|---|---|---|---|---|---|
| Nev | 1 | 1 | 11 | 203727.3 | 105805.4 |
| Nev | 1 | 2 | 16 | 183625 | 105754.7 |
| Nev | 2 | 3 | 16 | 135018.8 | 94628.04 |
| Nev | 2 | 4 | 13 | 156192.3 | 93304.72 |
| Nev | 2 | 5 | 20 | 192190 | 100400.5 |
| Nev | 3 | 6 | 12 | 151125 | 88063.07 |
| Vir | 4 | 7 | 13 | 197307.7 | 80288.13 |
| Vir | 4 | 8 | 14 | 168700 | 86626.27 |
| Vir | 5 | 9 | 6 | 178716.7 | 107857.3 |
| Vir | 5 | 10 | 9 | 159511.1 | 132957.2 |
| Vir | 5 | 11 | 9 | 150488.9 | 70977.03 |
| Vir | 6 | 12 | 11 | 212963.6 | 112784.7 |

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

This format is especially useful for creating a database containing only summary information such as the means, standard deviations, etc. To create a summary database, take the following steps:

1. Run this report on the data, summarizing across the categorical variables of interest.
2. Copy the output report to the clipboard.
3. Open a new database (or spreadsheet).
4. Paste the data from the clipboard to this new database by placing the cursor in the upper-left cell and pasting. The paste can use the Ctrl-V key or Paste from the Edit menu.
5. Label the columns in the Variable Info sheet.

Chapter 205

# T-Test – One-Sample or Paired

## Introduction

The procedure is used to compare the mean (or median) of a single group to a target value. To accomplish this, the procedure calculates the one-sample t-test, the paired t-test, the Wilcoxon Signed-Rank test, and the quantile (sign) test.

## Kinds of Research Questions

For the one-sample or paired-sample situation, the prime concern in research is examining a measure of central tendency (location) for the population of interest. The best-known measures of location are the mean and median. For a one-sample situation, we might want to know if the average waiting time in a doctor's office is greater than one hour, if the average refund on a 1040 tax return is different from $500, if the average assessment for similar residential properties is less than $120,000, or if the average growth of roses is 4 inches or more after two weeks of treatment with a certain fertilizer.

In the paired case, we take two measurements on the same individual at different times, or we have one measurement on each individual of a pair. Examples of the first case are two insurance-claim adjusters assessing the dollar damage for the same 15 cases or evaluation of the improvement in aerobic fitness for 15 subjects where measurements are made at the beginning of the fitness program and at the end of it. An example of the second paired situation is the testing of the effectiveness of two drugs, A and B, on 20 pairs of patients who have been matched on physiological and psychological variables. One patient in the pair receives drug A, and the other patient gets drug B.

The prime question relates to whether we have one random sample of observations or one random sample of pairs of observations. Given that determination, the second question focuses on whether the data are normally distributed. If normality is true, then the one-sample t-test is the choice for assessing whether the measure of central tendency, the mean, is different from some theoretical or hypothesized value. On the other hand, if normality is not valid, one of the two nonparametric tests, the Wilcoxon Signed Rank test or the quantile test, can be applied.

# Assumptions

This section describes the assumptions that are made when you use one of these tests. The key assumption relates to normality or nonnormality of the data. One of the reasons for the popularity of the t-test is its robustness in the face of assumption violation. However, if an assumption is not met even approximately, the significance levels and the power of the t-test are invalidated. Unfortunately, in practice it often happens that not one but several assumptions are not met. This makes matters even worse! Hence, take the steps to check the assumptions before you make important decisions based on these tests. Since the output includes items that let you investigate these assumptions, you should always do so.

## One-Sample T-Test Assumptions

The assumptions of the one-sample t-test are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution.
3. The sample is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

## Paired T-Test Assumptions

The assumptions of the paired t-test are:

1. The data are continuous (not discrete).
2. The data, i.e., the differences for the matched-pairs, follow a normal probability distribution.
3. The sample of pairs is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

## Wilcoxon Signed-Rank Test Assumptions

The assumptions of the Wilcoxon signed-rank test are as follows (note that the difference is between a data value and the hypothesized median or between the two data values of a pair):

1. The differences are continuous (not discrete).
2. The distribution of these differences is symmetric.
3. The differences are mutually independent.
4. The differences all have the same median.
5. The measurement scale is at least interval.

## Quantile Test Assumptions

The assumptions of the quantile (sign) test are:

1. A random sample has been taken resulting in observations that are independent and identically distributed.
2. The measurement scale is at least ordinal.

## Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can often ignore the continuous variable assumption. Perhaps the greatest restriction is that your data comes from a random sample of the population. If you do not have a random sample, your significance levels will definitely be incorrect.

## Bootstrapping

*Bootstrapping* was developed to provide standard errors and confidence intervals in situations in which the standard assumptions are not valid. In these nonstandard situations, bootstrapping is a viable alternative to the corrective action suggested earlier. The method is simple in concept, but it requires extensive computation time.

The bootstrap is simple to describe. You assume that your sample is actually the population and you draw *B* samples (*B* is over 1000) of N from the original dataset, with replacement. *With replacement* means that each observation may be selected more than once. For each bootstrap sample, the mean is computed and stored.

Suppose that you want the standard error and a confidence interval of the mean. The bootstrap sampling process has provided *B* estimates of the mean. The standard deviation of these *B* means is the bootstrap estimate of the standard error of the mean. The bootstrap confidence interval is found by arranging the *B* values in sorted order and selecting the appropriate percentiles from the list. For example, a 90% bootstrap confidence interval for the difference is given by fifth and ninety-fifth percentiles of the bootstrap mean values.

The main assumption made when using the bootstrap method is that your sample approximates the population fairly well. Because of this assumption, bootstrapping does not work well for small samples in which there is little likelihood that the sample is representative of the population. Bootstrapping should only be used in medium to large samples.

## Randomization Test

Because of the strict assumptions that must be made when using this procedure to test hypotheses about the difference, ***NCSS*** also includes a randomization test as outlined by Edgington (1987). Randomization tests are becoming more and more popular as the speed of computers allows them to be computed in seconds rather than hours.

A randomization test is conducted by enumerating all possible permutations of the signs of the values while leaving the data values in the original order. The mean is calculated for each permutation and the number of permutations that result in a mean with a magnitude greater than or equal to zero is counted. Dividing this count by the number of permutations tried gives the significance level of the test.

For even moderate sample sizes, the total number of permutations is in the trillions, so a Monte Carlo approach is used in which the permutations are found by random selection rather than complete enumeration. Edgington suggests that at least 1,000 permutations by selected. We suggest that this be increased to 10,000.

# Data Structure

In the one-sample case, there will be only one variable as shown for the variable Weight.

| Weight |
|---|
| 159 |
| 155 |
| 157 |
| 125 |
| 103 |
| 122 |
| 101 |
| 82 |
| 228 |
| 199 |
| 195 |
| 110 |
| 191 |
| 151 |
| 119 |
| 119 |
| 112 |
| 87 |
| 190 |
| 87 |
| 159 |
| 155 |
| 157 |

In the matched-pairs case, the analysis will require two variables. This example shows matched-pairs data with tire wear for the right and left tires of the same car.

| Right Tire | Left Tire |
|---|---|
| 42 | 54 |
| 75 | 73 |
| 24 | 22 |
| 56 | 59 |
| 52 | 51 |
| 56 | 45 |
| 23 | 29 |
| 55 | 58 |
| 46 | 49 |
| 52 | 58 |
| 47 | 49 |
| 62 | 67 |
| 55 | 58 |
| 62 | 64 |

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

These options specify the variables that will be used in the analysis. They also specify the type of analysis that will be performed. If you just specify Response Variables and leave Paired Variables blank, a One-Sample T-Test will be run. If you specify both a Response Variable and a Paired Variable, a Paired T-Test will be run comparing these two variables.

### Response Variables

#### Response Variable(s)

Specify one or more variables. If more than one variable is specified, a separate analysis is run for each variable.

## Paired Variables

### Paired Variable(s)

For paired measurements, the second variable is specified here. If this option is left blank, a One-Sample T-Test is run. If you specify a variable here, a Paired T-Test will be run. If multiple variables are specified in both Response Variable(s) and Paired Variable(s), the first variables in each list are compared, and then the second variables in each list are compared, and so on.

## Options

### H0 Value

The hypothesized value of the mean (or median for the nonparametric tests) if only one variable is specified. The hypothesized value of the mean (or median for the nonparametric tests) of the differences if two variables are specified. This value may represent a quantile other than the median if the Quantile Test Proportion is different from 0.5.

### Alpha Level

The value of alpha for the confidence limits, rejection decision, and power analysis. Usually, this number will range from 0.1 to 0.001. The default value of 0.05 results in 95% confidence limits.

### Quantile Test Proportion

This is the value of the binomial proportion used in the Quantile test. A value of 0.5 results in the Sign Test. Under the null hypothesis, the quantile test proportion is the proportion of all values below the null quantile.

## Resampling

### Bootstrap Confidence Intervals

This option causes bootstrap confidence intervals and all associated bootstrap reports and plots to be generated using resampling simulation. The bootstrap settings are set under the Resampling tab.

Bootstrapping may be time consuming when the bootstrap sample size is large. A reasonable strategy is to keep this option unchecked until you have considered all other reports. Then run this option with a bootstrap size of 100 and then 1000 to obtain an idea of the time needed to complete the simulation.

### Randomization Test

Check this option to run the randomization test.

Randomization tests may be time consuming when the Monte Carlo sample size is large. A reasonable strategy is to keep this option unchecked until you have run and considered all other reports. Then run this option with a Monte Carlo size of 100, then 1000, and then 10000 to obtain an idea of the time needed to complete the simulation.

# Reports Tab

The options on this panel control the format of the report.

## Select Additional Reports

### Nonparametric Tests

Select this option to display the indicated report.

## Select Plots

### Histogram … Average-Difference Plot

Check the boxes to display the plot.

## Report Options

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

# Histogram Tab

The options on this panel control the appearance of the histogram.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Histogram Settings

### Plot Style File

Designate a histogram style file. This file sets all histogram options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Histogram procedure.

### Number of Bars

Specify the number of intervals, bins, or bars used in the histogram.

## Titles

### Plot Title

This is the text of the title. The characters *{X}* are replaced by the name of the variable. Press the button on the right of the field to specify the font of the text.

# Probability Plot Tab

The options on this panel control the appearance of the probability plot.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Probability Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Probability Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

# Scatter Plot Tab

The options on this panel control the appearance of the scatter plot of the two paired variables.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the response variable. The characters *{X}* are replaced by the name of the paired variable. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Scatter Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

## Titles

### Plot Title

This is the text of the title. The following codes are replaced by appropriate values when the plot is generated.

{X} is replaced by the appropriate horizontal variable's name.

{Y} is replaced by the appropriate vertical variable's name.

{G} is replaced by the appropriate grouping variable's name.

{M} is replaced by the model (if available).

{S} is replaced by an appropriate internal phrase. This option works only for histograms.

{Z} is replaced by the appropriate variable's name (if used).

# Ave-Diff Plot Tab

The options on this panel control the appearance of the average-difference scatter plot.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by the appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Ave-Diff Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

## Titles

### Plot Title

This is the text of the title. The following codes are replaced by appropriate values when the plot is generated.

{X} is replaced by the appropriate horizontal variable's name.

{Y} is replaced by the appropriate vertical variable's name.

{G} is replaced by the appropriate grouping variable's name.

{M} is replaced by the model (if available).

{S} is replaced by an appropriate internal phrase. This option works only for histograms.

{Z} is replaced by the appropriate variable's name (if used).

# Resampling Tab

This panel controls the bootstrapping. Note that bootstrapping is only used when the Bootstrap report is checked on the Reports panel.

## Bootstrap Options – Sampling

### Samples (N)

This is the number of bootstrap samples used. A general rule of thumb is that you use at least 100 when standard errors are your focus or at least 1000 when confidence intervals are your focus. If computing time is available, it does not hurt to do 4000 or 5000.

We recommend setting this value to at least 3000.

### Retries

If the results from a bootstrap sample cannot be calculated, the sample is discarded and a new sample is drawn in its place. This parameter is the number of times that a new sample is drawn before the algorithm is terminated. We recommend setting the parameter to at least 50.

## Bootstrap Options – Estimation

### Percentile Type

The method used to create the percentiles when forming bootstrap confidence limits. You can read more about the various types of percentiles in the Descriptive Statistics chapter. We suggest you use the Ave *X*(p[n+1]) option.

### C.I. Method

This option specifies the method used to calculate the bootstrap confidence intervals. The reflection method is recommended.

- **Percentile**

  The confidence limits are the corresponding percentiles of the bootstrap values.

- **Reflection**

  The confidence limits are formed by reflecting the percentile limits. If *X*0 is the original value of the parameter estimate and *XL* and *XU* are the percentile confidence limits, the Reflection interval is (2 *X*0 - *XU*, 2 *X*0 - *XL*).

### Bootstrap Confidence Coefficients

These are the confidence coefficients of the bootstrap confidence intervals. Since bootstrapping calculations may take several minutes, it may be useful to obtain confidence intervals using several different confidence coefficients.

All values must be between 0.50 and 1.00. You may enter several values, separated by blanks or commas. A separate confidence interval is given for each value entered.

Examples:

0.90 0.95 0.99

0.90:.99(0.01)

0.90.

## Bootstrap Options – Histograms

### Vertical Axis Label

This is the label of the vertical axis of a bootstrap histogram.

### Horizontal Axis Label

This is the label of the horizontal axis of a bootstrap histogram.

### Plot Style File

This is the histogram style file. We have provided several different style files to choose from, or you can create your own in the Histogram procedure.

### Histogram Title

This is the title used on the bootstrap histograms.

### Number of Bars

The number of bars shown in a bootstrap histogram. We recommend setting this value to at least 25 when the number of bootstrap samples is over 1000.

## Randomization Test Options

### Monte Carlo Samples

Specify the number of Monte Carlo samples used when conducting randomization tests. You also need to check the 'Randomization Test' box under the Variables tab to run this test.

Somewhere between 1000 and 100000 Monte Carlo samples are usually necessary. Although the default is 1000, we suggest the use of 10000 when using this test.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running a Paired T-Test

This section presents an example of how to run a paired t-test. The data are the tire data shown above and found in the SAMPLE database. The data can be found under the variables labeled ***RtTire*** and ***LtTire***.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the T-Test – One Sample or Paired window.

**1 Open the SAMPLE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Sample.s0**.
- Click **Open**.

**2 Open the T-Test – One-Sample or Paired window.**

- On the menus, select **Analysis**, then **T-Tests**, then **T-Test - One Sample or Paired**. The T-Test - One Sample or Paired procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the T-Test - One Sample or Paired window, select the **Variables tab**. (This is the default.)
- Double-click in the **Response Variable(s)** text box. This will bring up the variable selection window.
- Select **RTTIRE** from the list of variables and then click **Ok**. "RTTIRE" will appear in the Response Variables box.
- Double-click in the **Paired Variable(s)** text box. This will bring up the variable selection window.
- Select **LTTIRE** from the list of variables and then click Ok. "LTTIRE" will appear in the Paired Variables box.
- Check the **Bootstrap Confidence Intervals** option.
- Check the **Randomization Test** option.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports and charts will be displayed in the Output window.

# Descriptive Statistics Section

| Variable | Count | Mean | Standard Deviation | Standard Error | 95% LCL of Mean | 95% UCL of Mean |
|---|---|---|---|---|---|---|
| RtTire | 14 | 50.5 | 13.96011 | 3.730996 | 42.43967 | 58.56033 |
| LtTire | 14 | 52.57143 | 13.7657 | 3.679038 | 44.62335 | 60.51951 |
| Difference | 14 | -2.071429 | 5.225151 | 1.39648 | -5.088341 | 0.9454835 |

T for Confidence Limits = 2.1604

### Variable

The name of the variable whose descriptive statistics are listed here. Note that the third row gives the statistics for the paired differences.

### Count

This is the number of nonmissing values.

### Mean

This is the average of the data values.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

### Standard Deviation

The sample deviation is the square root of the variance. It is a measure of dispersion based on squared distances from the mean for the variables listed.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

### Standard Error

This is the estimated standard deviation of the distribution of sample means for an infinite population.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

The standard error for the mean of differences is similar, except that *s* is computed on the differences themselves.

### Lower and Upper Confidence Limit

This formula gives the upper (with plus) and lower (with minus) values of a 100(1- ) interval estimate for the mean based on a t distribution with n-1 degrees of freedom. This interval estimate assumes that the population standard deviation is not known and that the data for this variable are normally distributed. This interval estimate is provided for the mean of the differences as well as for the mean of the two individual variables for paired data.

$$\bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}}$$

### T for Confidence Limits

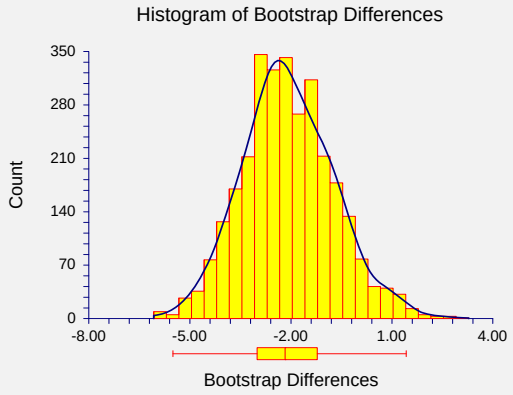
This is the value of $t_{\alpha/2,n-1}$ used to construct the above interval estimate.

# Bootstrap Section

| Estimation Results | | Bootstrap Confidence Limits | | |
|---|---|---|---|---|
| Parameter | Estimate | Conf. Level | Lower | Upper |
| **Mean** | | | | |
| Original Value | -2.0714 | 0.9000 | -4.4286 | 0.0714 |
| Bootstrap Mean | -2.0754 | 0.9500 | -5.0000 | 0.5000 |
| Bias (BM - OV) | -0.0040 | 0.9900 | -5.8571 | 1.1429 |
| Bias Corrected | -2.0675 | | | |
| Standard Error | 1.3590 | | | |

Sampling Method = Observation, Confidence Limit Type = Reflection, Number of Samples = 3000.

**Bootstrap Histograms Section**




This report provides bootstrap confidence intervals of the mean. Note that since these results are based on 3000 random bootstrap samples, they will differ slightly from the results you obtain when you run this report.

### Original Value

This is the parameter estimate obtained from the complete sample without bootstrapping.

### Bootstrap Mean

This is the average of the parameter estimates of the bootstrap samples.

### Bias (BM - OV)

This is an estimate of the bias in the original estimate. It is computed by subtracting the original value from the bootstrap mean.

### Bias Corrected

This is an estimated of the parameter that has been corrected for its bias. The correction is made by subtracting the estimated bias from the original parameter estimate.

### Standard Error

This is the bootstrap method's estimate of the standard error of the parameter estimate. It is simply the standard deviation of the parameter estimate computed from the bootstrap estimates.

### Conf. Level

This is the confidence coefficient of the bootstrap confidence interval given to the right.

### Bootstrap Confidence Limits – Lower and Upper

These are the limits of the bootstrap confidence interval with the confidence coefficient given to the left. These limits are computed using the confidence interval method (percentile or reflection) designated on the Bootstrap panel.

Note that to be accurate, these intervals must be based on over a thousand bootstrap samples and the original sample must be representative of the population.

### Bootstrap Histogram

The histogram shows the distribution of the bootstrap parameter estimates.

## Tests of Assumptions about Differences Section

**Tests of Assumptions about Differences Section**

| Assumption | Value | Probability | Decision(5%) |
|---|---|---|---|
| Skewness Normality | 1.3651 | 0.172212 | Cannot reject normality |
| Kurtosis Normality | 1.9065 | 0.056589 | Cannot reject normality |
| Omnibus Normality | 5.4982 | 0.063985 | Cannot reject normality |
| Correlation Coefficient | 0.929062 | | |

The main assumption when using the t-test is that the data are normally distributed. Either the single variable must be normal, or the differences for paired data must be normal. The normality assumption can be checked statistically by the skewness, kurtosis, or omnibus normality tests and visually by the normal probability plot or box plot.

In the case of nonnormality, the two nonparametric tests have the assumption of symmetry about the median. While the normal distribution is symmetric, not all symmetric distributions are normal. This assumption of symmetry is less restrictive than the one of normality, and it can be evaluated visually by the histogram or the normal probability plot. Generally, the Wilcoxon signed-rank test is more powerful than the sign test (and should be preferred), but there are some cases where the efficiency of the sign test surpasses that of the Wilcoxon signed-rank, specifically when the underlying distribution is a double exponential.

If the data are asymmetrical, the natural tendency is to use the nonparametric test. However, frequently a transformation, such as the natural logarithm or the square root of the original data, can change the underlying distribution from skewed to normal. To evaluate whether the underlying distribution of the variable is normal after the transformation, rerun the normal probability plot on the transformed variable. If some of the data values are negative or zero, it may be necessary to add a constant to the original data prior to the transformation. Of course, if the transformation or re-expression works, then the one-sample t-test is performed on the transformed data.

### Normality (Skewness, Kurtosis, and Omnibus)

These three tests allow you to test the skewness, kurtosis, and overall normality of the data. If any of them reject the hypothesis of normality, the data should not be considered normal. These tests are discussed in more detail in the Descriptive Statistics chapter.

# T-Test Section

**T-Test For Difference Between Means Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 | Power (Alpha=.05) | Power (Alpha=.01) |
|---|---|---|---|---|---|
| RtTire-LtTire<>0 | -1.4833 | .161824 | No | .279644 | .101545 |
| Randomization Test | | .146000 | No | | |
| RtTire-LtTire<0 | -1.4833 | .080912 | No | .405551 | .160410 |
| RtTire-LtTire>0 | -1.4833 | .919088 | No | .001124 | .000120 |

## Alternative Hypothesis

In hypothesis testing, the null and alternative hypotheses are always the opposite of one another. For instance, in a two-tail test on the difference between two paired means, the null hypothesis would be $H_o$: $\mu_d = 0$ with the alternative being $H_a$: $\mu_d \neq 0$. This two-tail alternative is represented by RtTire-LtTire<>0. The left-tail alternative is represented by RtTire-LtTire<0 (i.e., $H_a$: $\mu_d < 0$) while the right-tail alternative is depicted by RtTire-LtTire>0 (i.e., $H_a$: $\mu_d > 0$).

## T-Value

This is the test statistic for the t-test. It has n-1 degrees of freedom. It is identical for both one-tailed and two-tailed tests.

$$t_{n-1} = \frac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}}$$

## Prob Level

This is the significance level (or p-value) of the statistical test. It is the probability that the test statistic may take on a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. If the significance level is less than $\alpha$, say 5%, the null hypothesis is rejected. If the significance level is greater than $\alpha$, we do not have enough evidence to reject the null hypothesis.

Note that if a randomization test was selected, its probability level is displayed on the second line.

## Reject H0 at .050

This is the conclusion reached about the null hypothesis. It will be either ‘Yes’ or ‘No’ for a 5% level of significance. Note that when we say No, we really mean that we do not have enough evidence to reject H0. This is very different from concluding that the null hypothesis is true!

## Power(Alpha=0.05, Alpha=0.01)

Power is the probability of rejecting the null hypothesis when the alternative hypothesis is true. The power of a test is one minus the probability of a type II error ($\beta$). The power of a test depends on the value of the type I error, the sample size, the standard deviation, and the magnitude of the difference between the null and alternative hypothesized means. To calculate the power here, we set this difference to the actual difference observed in the sample.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of sensitivity in hypothesis testing. This estimate of power is based upon the sampling distribution of the statistic being normal under the alternative hypothesis.

## Nonparametric Tests Section

Nonparametric methods are also called distribution-free methods because they do not depend on a complete specification of the distribution shape. When the data are not normal, there are two possibilities: the quantile test and the Wilcoxon signed-rank test.

### Quantile (Sign) Test

The quantile (sign) test is perhaps the oldest of all the nonparametric procedures. This nonparametric test is based on the binomial distribution. It assumes two mutually exclusive outcomes, constant or stable probability of success or failure, and *n* independent trials. Some quantiles of interest are median, quartile, decile, and percentile.

When the quantile of interest is the median, a quantile test is called the *sign test*. The terminology, sign test, reinforces the point that the data are converted to a series of pluses and minuses. The test is based on the number of pluses that occur. Zero differences are thrown out, and the sample size is reduced accordingly.

While the sign test is simple, there are more powerful nonparametric alternatives, such as the Wilcoxon signed-rank test. However, if the shape of the underlying distribution of a variable is the double exponential distribution, the sign test may be the better choice.

**Quantile (Sign) Test**

| Null Quantile (Q0) | Quantile Proportion | Number Lower | Number Higher | H1:Q<>Q0 Prob Level | H1:Q<Q0 Prob Level | H1:Q>Q0 Prob Level |
|---|---|---|---|---|---|---|
| 0 | 0.5 | 10 | 4 | 0.179565 | 0.089783 | 0.971313 |

### Null Quantile (Q0)

Under the null hypothesis, the proportion of all values below the null quantile is the quantile proportion. For the sign test, the null quantile is the null median. For a paired sign test, the null quantile is often set to 0.

### Quantile Proportion

Under the null hypothesis, the quantile proportion is the proportion of all values below the null quantile. For the sign test, this proportion is 0.5.

### Number Lower

This is the actual number of values (or differences in a paired test) that are below the null quantile.

### Number Higher

This is the actual number of values (or differences in a paired test) that are above the null quantile.

### H1:Q<>Q0 Prob Level

This is the two-sided probability that the true quantile is equal to the stated null quantile (Q0), for the quantile proportion stated and given the observed values. A small prob level indicates that the true quantile for the stated quantile proportion is different from the null quantile.

### H1:Q<Q0 Prob Level

This is the one-sided probability that the true quantile is greater than or equal to the stated null quantile (Q0), for the quantile proportion stated and given the observed values. A small prob level indicates that the true quantile for the stated quantile proportion is less than the null quantile.

### H1:Q>Q0 Prob Level

This is the one-sided probability that the true quantile is less than or equal to the stated null quantile (Q0), for the quantile proportion stated and given the observed values. A small prob level indicates that the true quantile for the stated quantile proportion is greater than the null quantile.

### Wilcoxon Signed-Rank Test

This nonparametric test makes use of the sign and the magnitude of the rank of the differences (original data minus the hypothesized value for one-sample data or differences between the pairs of measurements for paired data). It is the best nonparametric alternative to the one sample t-test or paired t-test.

**Nonparametric Tests Section**

**Wilcoxon Signed-Rank Test for Difference in Medians**

| W Sum Ranks | Mean of W | Std Dev of W | Number of Zeros | Number Sets of Ties | Multiplicity Factor |
|---|---|---|---|---|---|
| 21 | 52.5 | 15.84692 | 0 | 3 | 126 |

| | Exact Probability | | Approximation Without Continuity Correction | | | Approximation With Continuity Correction | | |
|---|---|---|---|---|---|---|---|---|
| Alternative Hypothesis | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 |
| X1-X2<>0 | .049438 | Yes | 1.9878 | .046837 | Yes | 1.9562 | .050440 | No |
| X1-X2<0 | .024719 | Yes | -1.9878 | .023419 | Yes | -1.9562 | .025220 | Yes |
| X1-X2>0 | .979065 | No | -1.9878 | .976581 | No | -2.0193 | .978273 | No |

### Sum Ranks (W)

The basic statistic for this test is the sum of the positive ranks, $\Sigma R_+$ (The sum of the positive ranks is chosen arbitrarily. The sum of the negative ranks could equally be used). This statistic is called *W*.

$$W = \sum R_+$$

### Mean of W

This is the mean of the sampling distribution of the sum of ranks for a sample of *n* items.

$$\mu_W = \frac{n(n+1) - d_0(d_0+1)}{4}$$

where $d_0$ is the number of zero differences.

### Std Dev of W

This is the standard deviation of the sampling distribution of the sum of ranks. Here $t_i$ represents the number of times the $i^{th}$ value occurs.

$$s_W = \sqrt{\frac{n(n+1)(2n+1) - d_0(d_0+1)(2d_0+1)}{24} - \frac{\sum t_i^3 - \sum t_i}{48}}$$

where $d_0$ is the number zero differences, $t_i$ is the number of absolute differences that are tied for a given non-zero rank, and the sum is over all sets of tied ranks.

### Number of Zeros

This is the number of times that the difference between the observed value (or difference) and the hypothesized value is zero. The zeros are used in computing ranks, but are not considered positive ranks or negative ranks.

### Number Sets of Ties

The treatment of ties is to assign an average rank for the particular set of ties. This is the number of sets of ties that occur in the data, including ties at zero.

### Multiplicity Factor

This is the correction factor that appeared in the standard deviation of the sum of ranks when there were ties.

### Alternative Hypothesis

For the Wilcoxon signed-rank test, the null and alternative hypotheses relate to the median. In the two-tail test for the median difference (assuming a hypothesized value of 0), the null hypothesis would be $H_o$: median=0 with the alternative being $H_a$: median≠0. This two-tail alternative is represented by Median<>0.

The left-tail alternative is represented by Median<0 (i.e., $H_a$: median<0) while the right-tail alternative is depicted by Median>0 (i.e., $H_a$: median>0). For paired measurements, the hypothesized median is set equal to zero. If a value other than zero is desired for paired data, create a new single variable equal to the differences and rerun this test.

### Exact Probability: Prob Level

This is an exact p-value for this statistical test, assuming no ties. The p-value is the probability that the test statistic will take on a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. If the p-value is less than $\alpha$, say 5%, the null hypothesis is rejected. If the p-value is greater than $\alpha$, the null hypothesis is accepted. For convenience, the p-value is given for all three alternatives although only one is actually used.

### Exact Probability: Reject H0 at .050

This is the conclusion reached about the null hypothesis. It will be to either accept $H_o$ or reject $H_o$ at the assigned level of significance. An acceptance means that the null hypothesis is tenable, and a rejection means that it is not.

### Approximations with (and without) Continuity Correction: Z-Value

Given the sample size is at least ten, a normal approximation method may be used to approximate the distribution of the sum of ranks. Although this method does correct for ties, it does not have the continuity correction factor. The *z* value is as follows:

$$z = \frac{W - \mu_w}{\sigma_W}$$

If the correction factor for continuity is used, the formula becomes:

$$z = \frac{W - \mu_w \pm \frac{1}{2}}{\sigma_W}$$

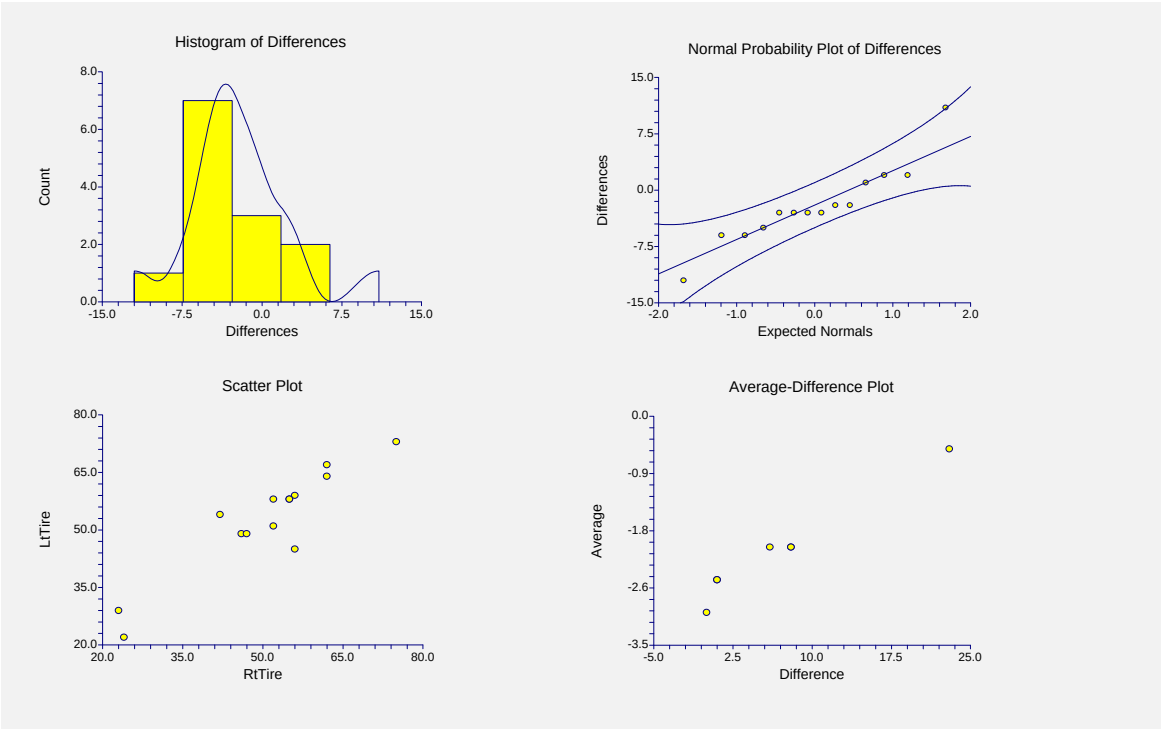
### Approximations with (and without) Continuity Correction: Prob Level

This is the p-value for the normal approximation approach for the Wilcoxon signed-rank test. The p-value is the probability that the test statistic will take a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. If the p-value is less than a, say 5%, the null hypothesis is rejected. If the p-value is greater than a, the null hypothesis is accepted.

### Approximations with (and without) Continuity Correction: Reject H0 at .050

This is the conclusion reached about the whether to reject null hypothesis. It will be either Yes or No at the given level of significance.

# Graphic Perspectives




## Histogram and Density Trace

The nonparametric tests need the assumption of symmetry, and these two graphic tools can provide that information. Since the histogram's shape is impacted by the number of classes or bins and the width of the bins, the best choice is to trust the density trace, which is a smoothed histogram. If the distribution of differences is symmetrical but not normal, proceed with the nonparametric test.

## Normal Probability Plot

If any of the observations fall outside the confidence bands, the data are not normal. The goodness-of-fit tests mentioned earlier, especially the omnibus test, should confirm this fact statistically. If only one observation falls outside the confidence bands and the remaining observations hug the straight line, there may be an outlier. If the data were normal, we would see the points falling along a straight line.

Note that these confidence bands are based on large-sample formulas. They may not be accurate for small samples.

## Scatter Plot

The intention of this plot is to look for patterns between the pairs. Preferably, you would like to see either no correlation or a positive linear correlation between Y and X. If there is a curvilinear relationship between Y and X, the paired t-test is not appropriate. If there is a negative relationship between the observations in the pairs, the paired t-test is not appropriate. If there are outliers, the nonparametric approach would be safer.

## Average-Difference Plot

This average-difference plot is designed to detect a lack of symmetry in the data. This plot is constructed from the paired differences, not the original data. Here's how. Let D(i) represent the ith ordered difference. Pairs of these sorted differences are considered, with the pairing being

done as you move toward the middle from either end. That is, consider the pairs D(1) and D(n), D(2) and D(n-1), D(3) and D(n-2), etc. Plot the average versus the difference of each of these pairs. Your plot will have about n/2 points, depending on whether n is odd or even. If the data are symmetric, the average of each pair will be the median and the difference between each pair will be zero.

Symmetry is an important assumption for the t-test. A perfectly symmetric set of data should show a vertical line of points hitting the horizontal axis at the value of the median. Departures from symmetry would deviate from this standard.

# One-Sample T-Test Checklist

This checklist, prepared by a professional statistician, is a flowchart of the steps you should complete to conduct a valid one-sample or paired-sample t-test (or one of its nonparametric counterparts). You should complete these tasks in order.

## Step 1 – Data Preparation

### Introduction

This step involves scanning your data for anomalies, data entry errors, typos, and so on. Frequently we hear of people who completed an analysis with the right techniques but obtained strange conclusions because they had mistakenly selected the data.

### Sample Size

The sample size (number of nonmissing rows) has a lot of ramifications. The larger the sample size for the one-sample t-test the better. Of course, the t-test may be performed on very small samples, say 4 or 5 observations, but it is impossible to assess the validity of assumptions with such small samples. It is our statistical experience that at least 20 observations are necessary to evaluate normality properly. On the other hand, since skewness can have unpleasant effects on t-tests with small samples, particularly for one-tailed tests, larger sample sizes (30 to 50) may be necessary.

It is possible to have a sample size that is too large for a statistical significance test. When your sample size is very large, you are almost guaranteed to find statistical significance. However, the question that then arises is whether the magnitude of the difference is of practical importance.

### Missing Values

The number and pattern of missing values is always an issue to consider. Usually, we assume that missing values occur at random throughout your data. If this is not true, your results will be biased since a particular segment of the population is underrepresented. If you have a lot of missing values, some researchers recommend comparing other variables with respect to missing versus nonmissing. If you find large differences in other variables, you should begin to worry about whether the missing values might cause a systematic bias in your results.

### Type of Data

The mathematical basis of the t-test assumes that the data are continuous. Because of the rounding that occurs when data are recorded, all data are technically discrete. The validity of assuming the continuity of the data then comes down to determining when we have too much

rounding. For example, most statisticians would not worry about human-age data that was rounded to the nearest year. However, if these data were rounded to the nearest ten years or further to only three groups (young, adolescent, and adult), most statisticians would question the validity of the probability statements. Some studies have shown that the t-test is reasonably accurate when the data has only five possible values (most would call this discrete data). If your data contains less than five unique values, any probability statements made are tenuous.

### Outliers

Generally, outliers cause distortion in statistical tests. You must scan your data for outliers (the box plot is an excellent tool for doing this). If you have outliers, you have to decide if they are one-time occurrences or if they would occur in another sample. If they are one-time occurrences, you can remove them and proceed. If you know they represent a certain segment of the population, you have to decide between biasing your results (by removing them) or using a nonparametric test that can deal with them. Most would choose the nonparametric test.

## Step 2 – Setup and Run the Panel

### Introduction

Now comes the fun part: running the program. ***NCSS*** is designed to be simple to operate, but it can still seem complicated. When you go to run a procedure such as this for the first time, take a few minutes to read through the chapter again and familiarize yourself with the issues involved.

### Enter Variables

The ***NCSS*** panels are set with ready-to-run defaults. About all you have to do is select the appropriate variable (variables for paired data).

### Select All Plots

As a rule, you should select all diagnostic plots (box plots, histograms, etc.) even though they may take a few extra seconds to generate. They add a great deal to your analysis of the data.

### Specify Alpha

Most beginners in statistics forget this important step and let the alpha value default to the standard 0.05. You should make a conscious decision as to what value of alpha is appropriate for your study. The 0.05 default came about when people had to rely on printed probability tables in which there were only two values available: 0.05 or 0.01. Now you can set the value to whatever is appropriate.

## Step 3 – Check Assumptions

### Introduction

Once the program output is displayed, you will be tempted to go directly to the probability of the t-test, determine if you have a significant result, and proceed to something else. However, it is very important that you proceed through the output in an orderly fashion. The first task is to determine which of the assumptions are met by your data.

Sometimes, when the data are nonnormal, a data transformation (like square roots or logs) might normalize the data. Frequently, this kind of transformation or re-expression approach works very well. However, always check the transformed variable to see if it is normally distributed.

It is not unusual in practice to find a variety of tests being run on the same basic null hypothesis. That is, the researcher who fails to reject the null hypothesis with the first test will sometimes try several others and stop when the hoped-for significance is obtained. For instance, a statistician might run the one-sample t-test on the original data, the one-sample t-test on the logarithmically transformed data, the Wilcoxon rank-sum test, and the Quantile test. An article by Gans (1984) suggests that there is no harm on the true significance level if no more than two tests are run. This is not a bad option in the case of questionable outliers. However, as a rule of thumb, it seems more honest to investigate whether the data is normal. The conclusion from that investigation should direct you to the right test.

### Random Sample

The validity of this assumption depends on the method used to select the sample. If the method used ensures that each individual in the population of interest has an equal probability of being selected for this sample, you have a random sample. Unfortunately, you cannot tell if a sample is random by looking at either it or statistics from it.

### Check Descriptive Statistics

You should check the Descriptive Statistics Section first to determine if the Count and the Mean are reasonable. If you have selected the wrong variable, these values will alert you.

### Normality

To validate this assumption, you would first look at the plots. Outliers will show up on the box plots and the probability plots. Skewness, kurtosis, more than one mode, and a host of other problems will be obvious from the density trace on the histogram. After considering the plots, look at the Tests of Assumptions Section to get numerical confirmation of what you see in the plots. Remember that the power of these normality tests is directly related to the sample size, so when the normality assumption is accepted, double-check that your sample is large enough to give conclusive results (at least 20).

### Symmetry

The nonparametric tests need the assumption of symmetry. The easiest ways to evaluate this assumption are from the density trace on the histogram or from the average-difference plot.

## Step 4 – Choose the Appropriate Statistical Test

### Introduction

After understanding how your data fit the assumptions of the various one-sample tests, you are ready to determine which statistical procedures will be valid. You should select one of the following three situations based on the status of the normality.

### Normal Data

Use the T-Test Section for hypothesis testing and the Descriptive Statistics Section for interval estimation.

### Nonnormal and Asymmetrical Data

Try a transformation, such as the natural logarithm or the square root, on the original data since these transformations frequently change the underlying distribution from skewed to normal. If some of the data values are negative or zero, add a constant to the original data prior to the transformation. If the transformed data is now normal, use the T-Test Section for hypothesis testing and the Descriptive Statistics Section for interval estimation.

### Nonnormal and Symmetrical Data

Use the Wilcoxon Rank-Sum Test or the Quantile Test for hypothesis testing.

## Step 5 – Interpret Findings

### Introduction

You are now ready to conduct your test. Depending on the nature of your study, you should look at either of the following sections.

### Hypothesis Testing

Here you decide whether to use a two-tailed or one-tailed test. The two-tailed test is the standard. If the probability level is less than your chosen alpha level, reject the null hypothesis of equality to a specified mean (or median) and conclude that the mean is different. Your next task is to look at the mean itself to determine if the size of the difference is of practical interest.

### Confidence Limits

The confidence limits let you put bounds on the size of the mean (for one independent sample) or mean difference (for dependent samples). If these limits are narrow and close to your hypothesized value, you might determine that even though your results are statistically significant, there is no practical significance.

## Step 6 – Record Your Results

Finally, as you finish a test, take a moment to jot down your impressions. Explain what you did, why you did it, what conclusions you reached, which outliers you deleted, areas for further investigation, and so on. Since this is a technical process, your short-term memory will not retain these details for long. These notes will be worth their weight in gold when you come back to this study a few days later!

# Example of Paired T-Test Steps

This example will illustrate the use of one-sample tests for paired data. A new four-lane road is going through the west end of a major metropolitan area. About 150 residential properties will be affected by the road. A random sample of 15 properties was selected. These properties were evaluated by two different property assessors. We are interested in determining whether there is any difference in their assessment. The assessments are recorded in thousands of dollars and are shown in the table. The assessment values are represented by Value1 and Value2 for the two property assessors.

| Value1 | Value2 |
|---|---|
| 118.5 | 117.1 |
| 154.2 | 159.6 |
| 130.8 | 136.5 |
| 154.8 | 146.9 |
| 131.4 | 136.0 |
| 104.1 | 99.7 |
| 154.9 | 157.8 |
| 97.6 | 96.1 |
| 140.0 | 144.8 |
| 116.9 | 112.4 |
| 129.6 | 129.1 |
| 108.2 | 114.5 |
| 108.6 | 113.7 |
| 178.3 | 194.3 |
| 92.9 | 8 |

## Step 1 – Data Preparation

These data are paired measurements. The sample size is smaller than you would like, but it is 10% of the current population. There are no missing values, and the use of the dollar value makes the data continuous.

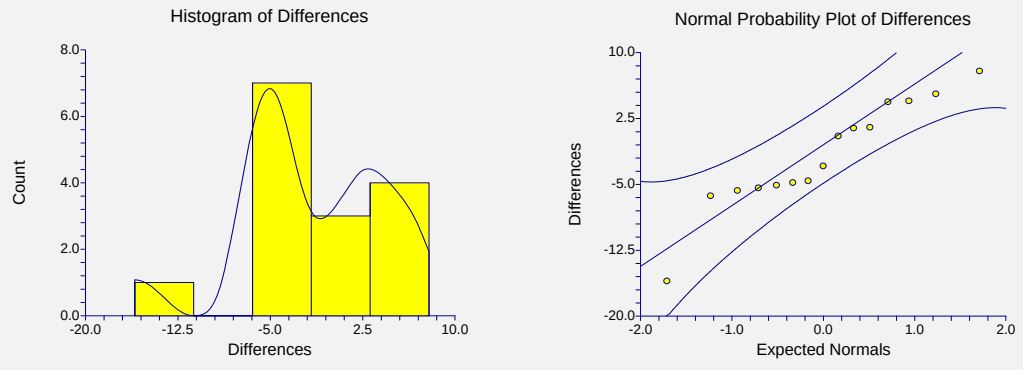
## Step 2 – Setup and Run the Paired T-Test Panel

The selection and running of the Paired T-Test from the Analysis menu on the pairs of assessments, Value1 and Value2, would produce the output that follows. The alpha value has been set at 0.05. Interpretation of the results will come in the steps to follow.

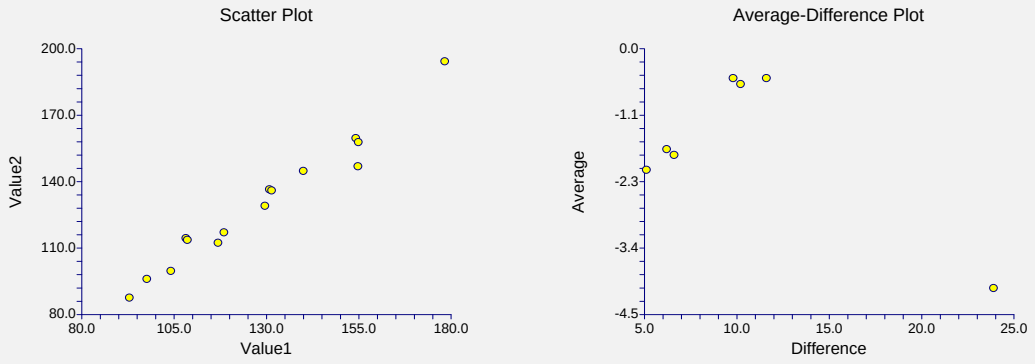
## Step 3 – Check Assumptions

The major assumption to check for is normality. We begin with the graphic perspectives: normal probability plots, histograms, density traces, and box plots. Since this is paired data, we look at the normality of the differences.

### Histogram, Density Trace, and Normal Probability Plot



## Scatter Plot and Average-Difference Plot




The normal probability plot on the differences indicates normality, except for an outlier on the low side. However, this potential outlier is within the 95% confidence bands of the probability plot. While the histogram and density trace are not good tools for evaluating normality on small samples, they do show the left skewness created by this one observation. This observation could be an outlier. Of course, a larger sample size would have been a definite advantage for the histogram and density trace, but normality seems to be valid (we make ourselves a note to check up on this outlier).

In evaluating normality by numerical measures, look at the Probability (p-value) and the Decision for the given alpha of 0.05. Investigation of the Tests of Assumptions Section confirms that the differences in assessment are normal by all three normality tests since the p-values are greater than 0.05. In fact, the p-values are much greater than 0.05. The “Cannot reject normality” under Decision(5%) is the formal conclusion of the normality tests.

**Tests of Assumptions Section**

| Assumption (About Differences) | Value | Probability | Decision(5%) |
|---|---|---|---|
| Skewness Normality | -0.9490 | 0.342635 | Cannot reject normality |
| Kurtosis Normality | 0.7722 | 0.440019 | Cannot reject normality |
| Omnibus Normality | 1.4968 | 0.473127 | Cannot reject normality |
| Correlation Coefficient | 0.982357 | | |

From the scatter plot above, it is evident that there is a strong positive linear relationship between the two assessments, as also confirmed by the Pearson correlation of 0.9824.

# Step 4 – Choose the Appropriate Statistical Test

In Step 3, the conclusions from checking the assumptions were three-fold: (1) the data are continuous, (2) the differences are normally distributed, and (3) there is a strong positive relationship between the two assessments. As a result of these findings, the appropriate statistical test is the paired t-test, which is shown next.

**Descriptive Statistics Section**

| Variable | Count | Mean | Standard Deviation | Standard Error | 95% LCL of Mean | 95% UCL of Mean |
|---|---|---|---|---|---|---|
| Value1 | 15 | 128.0533 | 24.68883 | 6.374629 | 114.3811 | 141.7256 |
| Value2 | 15 | 129.74 | 28.30113 | 7.307321 | 114.0674 | 145.4126 |
| Difference | 15 | -1.686667 | 6.140366 | 1.585436 | -5.087088 | 1.713755 |

T for Confidence Limits = 2.1448

**T-Test For Difference Between Means Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 | Power (Alpha=.05) | Power (Alpha=.01) |
|---|---|---|---|---|---|
| Value1-Value2<>0 | -1.0639 | .305402 | No | .168139 | .051619 |
| Value1-Value2<0 | -1.0639 | .152701 | No | .263633 | .086687 |
| Value1-Value2>0 | -1.0639 | .847299 | No | .003912 | .000489 |

## Step 5 – Interpret Findings

In the Descriptive Statistics Section, the mean difference is -$1.687 thousand with the standard deviation of differences being $6.140 thousand. The 95% interval estimate for the mean difference ranges from -$5.087 thousand to $1.714 thousand.

The formal two-tail hypothesis test for this example is shown under the T-Test Section. The p-value for this two-tail test is 0.305402, which is much greater than 0.05. Thus, the conclusion of this hypothesis test is acceptance, i.e., there is no difference in the assessments. However, it is important to note that the power of this test is only 0.168139. One would like the power to be at least .80 or more, but small sample sizes will have poor power unless the difference is very pronounced.

Remember when checking the assumption of normality, we noted that there was one possible outlier in the normal probability plot in the output. If we had run the Wilcoxon Signed-Rank test instead of the paired t-test, the p-value would be 0.302795. Hence, the conclusion is the same: there is no difference between assessments. This kind of decision confirmation does not always happen, but it is a simple option on questionable assumption situations. However, since the data are normally distributed, the paired t-test was the correct statistical test to choose.

**Wilcoxon Signed-Rank Test for Difference in Medians**

| W Sum Ranks | Mean of W | Std Dev of W | Number of Zeros | Number Sets of Ties | Multiplicity Factor |
|---|---|---|---|---|---|
| 41 | 60 | 17.60682 | 0 | 0 | 0 |

| | Exact Probability | | Approximation Without Continuity Correction | | | Approximation With Continuity Correction | | |
|---|---|---|---|---|---|---|---|---|
| Alternative Hypothesis | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 |
| X1-X2<>0 | .302795 | No | 1.0791 | .280531 | No | 1.0507 | .293383 | No |
| X1-X2<0 | .151398 | No | -1.0791 | .140265 | No | -1.0507 | .146691 | No |
| X1-X2>0 | .861572 | No | -1.0791 | .859735 | No | -1.1075 | .865967 | No |

## Step 6 – Record Your Results

The conclusions for this example are that there is no difference between assessors for residential properties evaluated in this area, according to the paired t-test. The Wilcoxon Signed-Rank gave the same conclusion. If you were troubled by the one outlier, you could use a transformation on the differences plus a constant and rerun the paired t-test. Or, further examination of the one outlier might reveal extenuating circumstances that confirm that this is a one-time anomaly. If that were the case, the observation could be omitted and the analysis redone.

# Example of One-Sample T-Test Steps

This example will illustrate the use of one-sample tests for a single variable. A registration service for a national motel/hotel chain wants the average wait time for incoming calls on Mondays (during normal business hours, 8:00 a.m. to 5:00 p.m.) to be less than 25 seconds. A random sample of 30 calls yielded the results shown below.

| Row | Anstime | Row | Anstime |
|---|---|---|---|
| 1 | 15 | 16 | 8 |
| 2 | 12 | 17 | 12 |
| 3 | 25 | 18 | 30 |
| 4 | 11 | 19 | 12 |
| 5 | 20 | 20 | 25 |
| 6 | 10 | 21 | 26 |
| 7 | 16 | 22 | 16 |
| 8 | 26 | 23 | 29 |
| 9 | 21 | 24 | 22 |
| 10 | 23 | 25 | 12 |
| 11 | 32 | 26 | 12 |
| 12 | 34 | 27 | 12 |
| 13 | 17 | 28 | 30 |
| 14 | 16 | 29 | 15 |
| 15 | 16 | 30 | 39 |

## Step 1 – Data Preparation

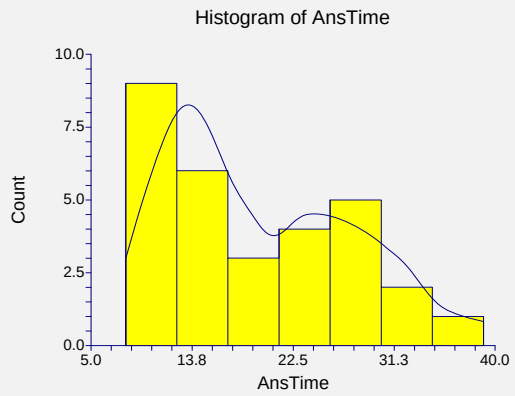
This is not paired data but just a single random sample of one variable. There are no missing values, and the variable is continuous.

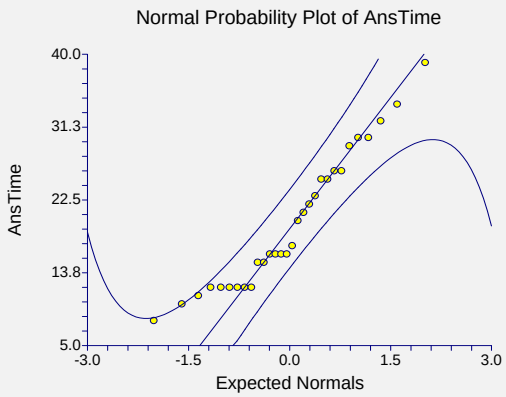
## Step 2 – Setup and Run the One-Sample T-Test

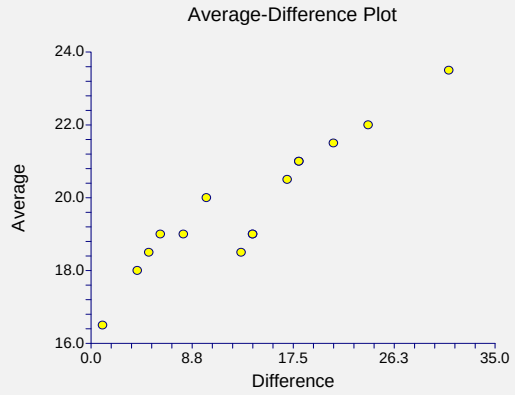
Select and run the One-Sample T-Test from the Analysis menu on the single variable, Anstime. The alpha value has been set at 0.05. Interpretation of the results will come in the steps to follow.

## Step 3 – Check Assumptions

The major assumption to check for is normality, and you should begin with the graphic perspectives: normal probability plots, histograms, density traces, and box plots. Some of these plots are given below.




The normal probability plot above does not look straight. It shows some skewness to the right. Some of the data points fall outside the 95% confidence bands. The histogram and density trace on answer time confirm the skewness to the right. This type of skewness to the right turns up quite often when dealing with elapsed-time data.

The skewness, kurtosis, and the omnibus normality tests in the output below have p-values greater than 0.05, indicating that answer time seems to be normally distributed. This conflict in conclusions between the normal probability plot and the normality tests is probably due to the fact that this sample size is not large enough to accurately assess the normality of the data.

**Tests of Assumptions Section**

| Assumption | Value | Probability | Decision(5%) |
|---|---|---|---|
| Skewness Normality | 1.4246 | 0.154281 | Cannot reject normality |
| Kurtosis Normality | -0.7398 | 0.459446 | Cannot reject normality |
| Omnibus Normality | 2.5766 | 0.275733 | Cannot reject normality |

## Step 4 – Choose the Appropriate Statistical Test

In Step 3, the conclusions from checking the assumptions were two-fold: (1) the data are continuous, and (2) the answer times are (based on the probability plot) non-normal. As a result of these findings, the appropriate statistical test is the Wilcoxon Signed-Rank test, which is shown in the figure. For comparison purposes, the t-test results are also shown in the output.

**T-Test For Difference Between Mean and Value Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 | Power (Alpha=.05) | Power (Alpha=.01) |
|---|---|---|---|---|---|
| AnsTime<>25 | -3.4744 | .001630 | Yes | .918832 | .757472 |
| AnsTime<25 | -3.4744 | .000815 | Yes | .959664 | .837398 |
| AnsTime>25 | -3.4744 | .999185 | No | .000000 | .000000 |

**Wilcoxon Signed-Rank Test**

| W Sum Ranks | Mean of W | Std Dev of W | Number of Zeros | Number Sets of Ties | Multiplicity Factor |
|---|---|---|---|---|---|
| 90 | 231 | 48.52319 | 2 | 8 | 384 |

| | Exact Probability | | Approximation Without Continuity Correction | | | Approximation With Continuity Correction | | |
|---|---|---|---|---|---|---|---|---|
| Alternative Hypothesis | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 |
| Median<>25 | | | 2.9058 | .003663 | Yes | 2.8955 | .003785 | Yes |
| Median<25 | | | -2.9058 | .001831 | Yes | -2.8955 | .001893 | Yes |
| Median>25 | | | -2.9058 | .998169 | No | -2.9161 | .998228 | No |

## Step 5 – Interpret Findings

Since the nonparametric test is more appropriate here and the concern was that the average answer time was less than 25 seconds, the Median<25 is the proper alternative hypothesis. The p-value for the Wilcoxon Signed-Rank test is 0.00128, which is much less than 0.05. Thus, the conclusion of the test is to reject the null hypothesis. This says that the median answer time is significantly less than 25 seconds.

It is interesting to note that the p-value for the left-tailed t-test is about the same. This points out the robustness of the t-test in the cases of heavy-tailed but almost symmetric distributions.

## Step 6 – Record Your Results

The conclusions for this example are that the median is less than 25 seconds. Again, if you were troubled by the shape of the distribution, you could use a transformation such as the natural logarithm to make the data more normal and try the t-test. However, in this case, that seems to be more work than is needed.

# Chapter 206

# T-Test – Two-Sample

## Introduction

This procedure calculates the two-sample t-test, the Mann-Whitney U test, and the Kolmogorov-Smirnov test of data either contained in two variables (columns) or in one variable indexed by a second (grouping) variable.

### Kinds of Research Questions

One of the most common tasks in research is to compare two populations (groups). We might want to compare the income level of two regions, the nitrogen content of two lakes, or the effectiveness of two drugs. The first question that arises is what aspects (parameters) of the populations we shall compare. We might consider comparing the averages, the medians, the standard deviations, the distributional shapes (histograms), or maximum values. We base the comparison parameter on our particular problem.

Perhaps the simplest comparison that we can make is between the means of the two populations. If we can show that the mean of population A is different from that of population B, we can conclude that the populations are different. Other aspects of the two populations can (and should) also be considered, but the mean is usually the starting point.

If we are willing to make assumptions about the other features of the two populations (such as that they are normally distributed and their variances are equal), we can use the two-sample t-test to compare the means of random samples drawn from these two populations. If these assumptions are violated, the nonparametric Mann-Whitney U test or the Kolmogorov-Smirnov test may be used instead.

## Assumptions

The following assumptions are made by the statistical tests described in this section. One of the reasons for the popularity of the t-test is its robustness in the face of assumption violation. However, if an assumption is not met even approximately, the significance levels and the power of the t-test are invalidated. Unfortunately, in practice it often happens that not one but several assumptions are not met. This makes matters even worse! Hence, take the appropriate steps to check the assumptions before you make important decisions based on these tests. Since the output includes items that let you investigate these assumptions, you should always do so.

### Two-Sample T-Test Assumptions

The assumptions of the two-sample t-test are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution.
3. The variances of the two populations are equal. (If not, the Aspin-Welch Unequal-Variance test is used.)
4. The two samples are independent. There is no relationship between the individuals in one sample as compared to the other (as there is in the paired t-test).
5. Both samples are simple random samples from their respective populations. Each individual in the population has an equal probability of being selected in the sample.

### Mann-Whitney U Test Assumptions

The assumptions of the Mann-Whitney U test are:

1. The variable of interest is continuous (not discrete). The measurement scale is at least ordinal.
2. The probability distributions of the two populations are identical, except for location.
3. The two samples are independent.
4. Both samples are simple random samples from their respective populations. Each individual in the population has an equal probability of being selected in the sample.

### Kolmogorov-Smirnov Test Assumptions

The assumptions of the Kolmogorov-Smirnov test are:

1. The measurement scale is at least ordinal.
2. The probability distributions are continuous.
3. The two samples are mutually independent.
4. Both samples are simple random samples from their respective populations.

## Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can often ignore the continuous variable assumption. Perhaps the greatest restriction is that your data come from a random sample of the population. If you do not have a random sample, your significance levels will definitely be incorrect.

## Bootstrapping

*Bootstrapping* was developed to provide standard errors and confidence intervals in situations in which the standard assumptions are not valid. In these nonstandard situations, bootstrapping is a viable alternative to the corrective action suggested earlier. The method is simple in concept, but it requires extensive computation time.

The bootstrap is simple to describe. You assume that your sample is actually the population and you draw *B* samples (*B* is over 1000) of N1 from the original group one dataset and N2 from the original group 2 dataset, with replacement. *With replacement sampling* means that each observation is placed back in the population before the next one is selected so that each observation may be selected more than once. For each bootstrap sample, the means and their difference are computed and stored.

Suppose that you want the standard error and a confidence interval of the difference. The bootstrap sampling process has provided *B* estimates of the difference. The standard deviation of these *B* differences is the bootstrap estimate of the standard error of the difference. The bootstrap confidence interval is found by arranging the *B* values in sorted order and selecting the appropriate percentiles from the list. For example, a 90% bootstrap confidence interval for the difference is given by fifth and ninety-fifth percentiles of the bootstrap difference values.

The main assumption made when using the bootstrap method is that your sample approximates the population fairly well. Because of this assumption, bootstrapping does not work well for small samples in which there is little likelihood that the sample is representative of the population. Bootstrapping should only be used in medium to large samples.

## Randomization Test

Because of the strict assumptions that must be made when using this procedure to test hypotheses about the difference, **NCSS** also includes a randomization test as outlined by Edgington (1987). Randomization tests are becoming more and more popular as the speed of computers allows them to be computed in seconds rather than hours.

A randomization test is conducted by enumerating all possible permutations of the groups while leaving the data values in the original order. The difference is calculated for each permutation and the number of permutations that result in a difference with a magnitude greater than or equal to the actual difference is counted. Dividing this count by the number of permutations tried gives the significance level of the test.

For even moderate sample sizes, the total number of permutations is in the trillions, so a Monte Carlo approach is used in which the permutations are found by random selection rather than complete enumeration. Edgington suggests that at least 1,000 permutations by selected. We suggest that this be increased to 10,000.

# Data Structure

The data may be entered in two formats, as shown in the two examples below. The examples give the yield of corn for two types of fertilizer. The first format is shown in the first table in which the responses for each group are entered in separate variables. That is, each variable contains all responses for a single group. In the second format the data are arranged so that all responses are entered in a single variable. A second variable, the Grouping Variable, contains an index that gives the group (A or B) to which the row of data belongs.

In most cases, the second format is more flexible. Unless there is some special reason to use the first format, we recommend that you use the second.

**Two Response Variables**

| Yield A | Yield B |
|---|---|
| 452 | 546 |
| 874 | 547 |
| 554 | 774 |
| 447 | 465 |
| 356 | 459 |
| 754 | 665 |
| 558 | 467 |
| 574 | 365 |
| 664 | 589 |
| 682 | 534 |
| | 456 |
| 547 | 651 |
| | 654 |
| 435 | 665 |
| | 546 |
| 245 | 537 |

**Grouping and Response Variables**

| Fertilizer | Yield |
|---|---|
| B | 546 |
| B | 547 |
| B | 774 |
| B | 465 |
| B | 459 |
| B | 665 |
| B | 467 |
| B | 365 |
| B | 589 |
| B | 534 |
| B | 456 |
| B | 651 |
| B | 654 |
| B | 665 |
| B | 546 |
| B | 537 |
| A | 452 |
| A | 874 |
| A | 554 |
| A | 447 |
| A | 356 |
| A | 754 |
| A | 558 |
| A | 574 |
| A | 664 |
| A | 682 |
| A | 547 |
| A | 435 |
| A | 245 |

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

Following is a list of the procedure's options.

## Variables Tab

The options on this panel specify which variables to use.

### Response Variables

#### Response Variable(s)

This option lets you specify the variable(s) to be analyzed. Note that if you specify only one variable here, you must also specify a grouping variable. If you simply want to compare two variables, you should specify them both here. Note that if more than one variable is specified, only the variable numbers are displayed.

### Group Variables

#### Group Variables

Optional group (breakdown) variables may be selected to indicate how the values of the response variable should be grouped. Examples of grouping variables are males and females, age groups, and yes or no responses. A separate analysis is performed for each pair of unique values in this variable. Note that the values in the variable can be either numeric or text. If more than one Group Variable is specified, a separate analysis is performed for all combinations of values.

### Options

#### H0 Value

This is the hypothesized difference between the two population means. It is usually assumed to be zero.

#### Alpha Level

The value of alpha for the confidence limits, rejection decision, and power analysis. Usually, this number will range from 0.1 to 0.001. The 0.05 default level represents 95% confidence limits.

### Resampling

#### Bootstrap Confidence Intervals

This option causes bootstrap confidence intervals and all associated bootstrap reports and plots to be generated using resampling simulation. The bootstrap settings are set under the Resampling tab.

Bootstrapping may be time consuming when the bootstrap sample size is large. A reasonable strategy is to keep this option unchecked until you have considered all other reports. Then run this option with a bootstrap size of 100 and then 1000 to obtain an idea of the time needed to complete the simulation.

### Randomization Test of Difference

Check this option to run the randomization test.

Randomization tests may be time consuming when the Monte Carlo sample size is large. A reasonable strategy is to keep this option unchecked until you have run and considered all other reports. Then run this option with a Monte Carlo size of 100, then 1000, and then 10000 to obtain an idea of the time needed to complete the simulation.

## Reports Tab

The options on this panel control the format of the report.

### Select Additional Reports

#### Nonparametric Tests

Select this option to display the indicated report.

### Select Plots

#### Histogram … Box Plot

Check the boxes to display the plot.

### Report Options

#### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

#### Value Labels

This option applies to the *Group Variable(s)*. It lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

#### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

# Histogram Tab

The options on this panel control the appearance of the histogram.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Histogram Settings

### Plot Style File

Designate a histogram style file. This file sets all histogram options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Histogram procedure.

### Number of Bars

Specify the number of intervals, bins, or bars used in the histogram.

## Titles

### Plot Title

This is the text of the title. The characters *{X}* are replaced by the name of the variable. Press the button on the right of the field to specify the font of the text.

# Probability Plot Tab

The options on this panel control the appearance of the probability plot.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on each axis. If left blank, these values are calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along each axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Probability Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Probability Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

# Box Plot Tab

The options on this panel control the appearance of the box plot.

## Vertical Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the axis. If left blank, these values are calculated from the data.

### Tick Label Settings…

Pressing this button brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along this axis.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on this axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* are replaced by the name of the variable. The characters *{M}* are replaced by the name of the selected probability distribution. Press the button on the right of the field to specify the font of the text.

### Tick Label Settings

Pressing this button brings up a window that sets the font, rotation, and number of decimal places displayed in the tick labels along this axis.

## Box Plot Settings

### Plot Style File

Designate a box plot style file. This file sets all box plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Box Plot procedure.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by the appropriate variable names. Press the button on the right of the field to specify the font of the text.

# Resampling Tab

This panel controls the bootstrapping. Note that bootstrapping is only used when the Bootstrap report is checked on the Reports panel.

## Bootstrap Options – Sampling

### Samples (N)

This is the number of bootstrap samples used. A general rule of thumb is that you use at least 100 when standard errors are your focus or at least 1000 when confidence intervals are your focus. If computing time is available, it does not hurt to do 4000 or 5000.

We recommend setting this value to at least 3000.

### Retries

If the results from a bootstrap sample cannot be calculated, the sample is discarded and a new sample is drawn in its place. This parameter is the number of times that a new sample is drawn before the algorithm is terminated. We recommend setting the parameter to at least 50.

## Bootstrap Options – Estimation

### Percentile Type

The method used to create the percentiles when forming bootstrap confidence limits. You can read more about the various types of percentiles in the Descriptive Statistics chapter. We suggest you use the Ave *X*(p[n+1]) option.

### C.I. Method

This option specifies the method used to calculate the bootstrap confidence intervals. The reflection method is recommended.

- **Percentile**

  The confidence limits are the corresponding percentiles of the bootstrap values.

- **Reflection**

  The confidence limits are formed by reflecting the percentile limits. If *X*0 is the original value of the parameter estimate and *XL* and *XU* are the percentile confidence limits, the Reflection interval is (2 *X*0 - *XU*, 2 *X*0 - *XL*).

### Bootstrap Confidence Coefficients

These are the confidence coefficients of the bootstrap confidence intervals. Since bootstrapping calculations may take several minutes, it may be useful to obtain confidence intervals using several different confidence coefficients.

All values must be between 0.50 and 1.00. You may enter several values, separated by blanks or commas. A separate confidence interval is given for each value entered.

Examples:

0.90 0.95 0.99

0.90:.99(0.01)

0.90.

## Bootstrap Options – Histograms

### Vertical Axis Label

This is the label of the vertical axis of a bootstrap histogram.

### Hortizontal Axis Label

This is the label of the horizontal axis of a bootstrap histogram.

### Plot Style File

This is the histogram style file. We have provided several different style files to choose from, or you can create your own in the Histogram procedure.

### Histogram Title

This is the title used on the bootstrap histograms.

### Number of Bars

The number of bars shown in a bootstrap histogram. We recommend setting this value to at least 25 when the number of bootstrap samples is over 1000.

## Randomization Test Options

### Monte Carlo Samples

Specify the number of Monte Carlo samples used when conducting randomization tests. You also need to check the 'Randomization tests' box under the Variables tab to run this test.

Somewhere between 1000 and 100000 Monte Carlo samples are usually necessary. Although the default is 1000, we suggest the use of 10000 when using this test.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running a Paired T-Test

This section presents an example of how to run a two-sample t-test. We will use the corn yield data found in YldA and YldB of the SAMPLE database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the T-Test – Two-Sample window.

**1 Open the SAMPLE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Sample.s0**.
- Click **Open**.

**2 Open the T-Test – Two-Sample window.**

- On the menus, select **Analysis**, then **T-Tests**, then **T-Test – Two-Sample**. The T-Test – Two-Sample procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the T-Test – Two-Sample window, select the **Variables tab**.
- Double-click in the **Response Variable(s)** box. This will bring up the variable selection window.
- Select **YldA** and **YldB** from the list of variables and then click Ok. The words "YldA-YldB" will appear in the Response Variables box.
- Check the **Bootstrap Confidence Intervals** option.
- Check the **Randomization Test of Difference** option.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports and charts will be displayed in the Output window.

## Descriptive Statistics Section

This section gives a descriptive summary of each group. See the Descriptive Statistics chapter for details about this section.

You should glance through this report quickly to make sure that the Means and Counts are correct. This provides another check of whether you are analyzing the data you intended!

| Variable | Count | Mean | Standard Deviation | Standard Error | 95% LCL of Mean | 95% UCL of Mean |
|---|---|---|---|---|---|---|
| YldA | 13 | 549.3846 | 168.7629 | 46.80641 | 447.4022 | 651.367 |
| YldB | 16 | 557.5 | 104.6219 | 26.15546 | 501.7509 | 613.249 |

Note: T-alpha (YldA) = 2.1788, T-alpha (YldB) = 2.1314

### Variable

These are the names of the independent variables.

### Count

The count gives the number of nonmissing values. This value is often referred to as the sample size or *n*.

### Mean

This is the average for each group.

### Standard Deviation

The sample standard deviation is the square root of the sample variance. It is a measure of spread.

### Standard Error

This is the estimated standard deviation for the distribution of sample means for an infinite population. It is the sample standard deviation divided by the square root of sample size, *n*.

### LCL of the Mean

This is the lower value of a 100(1-α)% interval estimate of the mean based on a Student's t distribution with n-1 degrees of freedom. This interval estimate assumes that the population standard deviation is not known and that the data are normally distributed.

### UCL of the Mean

This is the upper value of a 100(1-α)% interval estimate for the mean based on a t distribution with n-1 degrees of freedom.

### T-alpha

This is the t-value used to construct the confidence interval estimate. If you were constructing the interval manually, you would obtain this value from a table of the Student's t distribution.

## Confidence-Limits of Difference Section

Given that the assumptions of independent samples and normality are valid, this section provides an interval estimate (confidence limits) of the difference between the two means. Results are given for both the equal and unequal variance cases. Use the Equal Variance results if the Equal-Variance Test in the Tests of Assumptions Section is marked "Cannot reject." Otherwise, use the Unequal Variance results.

| Variance Assumption | DF | Mean Difference | Standard Deviation | Standard Error | 95% LCL of Difference | 95% UCL of Difference |
|---|---|---|---|---|---|---|
| Equal | 27 | -8.115385 | 136.891 | 51.11428 | -112.9932 | 96.76247 |
| Unequal | 19.1690 | -8.115385 | 198.5615 | 53.61855 | -120.2734 | 104.0426 |

Note: T-alpha (Equal) = 2.0518, T-alpha (Unequal) = 2.0918

### DF

The degrees of freedom for the two cases are next.

For the equal variance case:

$$df = n_X + n_Y - 2$$

For the unequal variance case:

$$df = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{s_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{s_Y^2}{n_Y}\right)^2}{n_Y - 1}}$$

### Mean Difference

This is the difference between the sample means.

### Standard Deviation

In the equal variance case, this quantity is:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{(n-1)s_X^2 + (n-1)s_Y^2}{n_X + n_Y - 2}}$$

In the unequal variance case, this quantity is:

$$s_{\bar{X}-\bar{Y}} = \sqrt{s_X^2 + s_Y^2}$$

### Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

For the equal variance case:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\left(\frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}\right)\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}$$

For the unequal variance case:

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}$$

### LCL of Difference

This is the lower value of a 100(1-α)% interval estimate for the difference between two means. The Equal Variance results are based on the usual t distribution. The Unequal Variance results are based on the Aspin-Welch Unequal-Variance procedure.

### UCL of Difference

This is the upper value of a 100(1-α)% interval estimate for the difference between two means. The Equal Variance results are based on the usual t distribution. The Unequal Variance results are based on the Aspin-Welch Unequal-Variance procedure.

### T-alpha

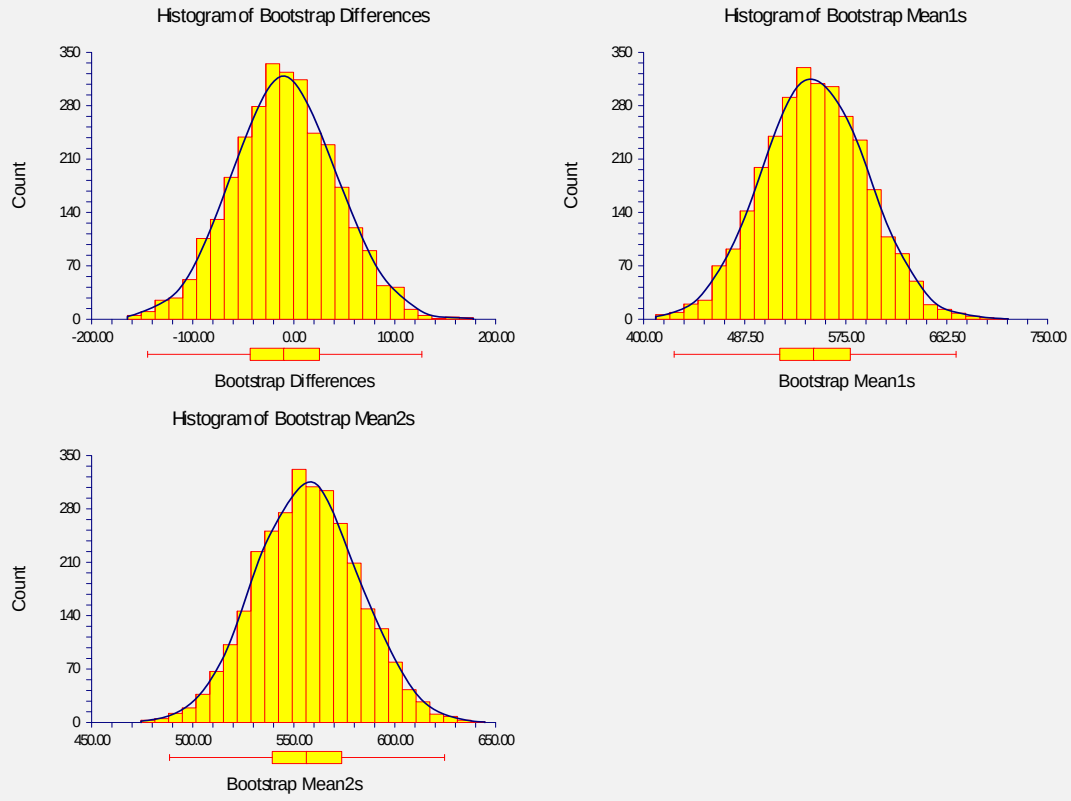
This is the t value used to construct the confidence limits. It depends upon the variance situation and the α level of significance (or 1-α degree of confidence).

# Bootstrap Section

| Estimation Results | | Bootstrap Confidence Limits | | |
|---|---|---|---|---|
| Parameter | Estimate | Conf. Level | Lower | Upper |
| **Mean Difference** | | | | |
| Original Value | -8.1154 | 0.9000 | -90.5036 | 75.0834 |
| Bootstrap Mean | -8.7702 | 0.9500 | -108.3319 | 91.0124 |
| Bias (BM - OV) | -0.6548 | 0.9900 | -135.1548 | 122.6257 |
| Bias Corrected | -7.4605 | | | |
| Standard Error | 50.7672 | | | |
| **Mean 1** | | | | |
| Original Value | 549.3846 | 0.9000 | 476.4615 | 624.9808 |
| Bootstrap Mean | 548.1727 | 0.9500 | 465.5423 | 635.9904 |
| Bias (BM - OV) | -1.2119 | 0.9900 | 430.7788 | 664.6100 |
| Bias Corrected | 550.5965 | | | |
| Standard Error | 44.6132 | | | |
| **Mean 2** | | | | |
| Original Value | 557.5000 | 0.9000 | 515.5688 | 598.9938 |
| Bootstrap Mean | 556.9429 | 0.9500 | 508.6281 | 606.5625 |
| Bias (BM - OV) | -0.5571 | 0.9900 | 492.3141 | 622.4988 |
| Bias Corrected | 558.0571 | | | |
| Standard Error | 25.1733 | | | |

Sampling Method = Observation, Confidence Limit Type = Reflection, Number of Samples = 3000.

**Bootstrap Histograms Section**




This report provides bootstrap confidence intervals of the two means and their difference. Note that since these results are based on 3000 random bootstrap samples, they will differ slightly from the results you obtain when you run this report.

### Original Value

This is the parameter estimate obtained from the complete sample without bootstrapping.

### Bootstrap Mean

This is the average of the parameter estimates of the bootstrap samples.

### Bias (BM - OV)

This is an estimate of the bias in the original estimate. It is computed by subtracting the original value from the bootstrap mean.

### Bias Corrected

This is an estimated of the parameter that has been corrected for its bias. The correction is made by subtracting the estimated bias from the original parameter estimate.

### Standard Error

This is the bootstrap method's estimate of the standard error of the parameter estimate. It is simply the standard deviation of the parameter estimate computed from the bootstrap estimates.

### Conf. Level

This is the confidence coefficient of the bootstrap confidence interval given to the right.

### Bootstrap Confidence Limits - Lower and Upper

These are the limits of the bootstrap confidence interval with the confidence coefficient given to the left. These limits are computed using the confidence interval method (percentile or reflection) designated on the Bootstrap panel.

Note that to be accurate, these intervals must be based on over a thousand bootstrap samples and the original sample must be representative of the population.

### Bootstrap Histogram

The histogram shows the distribution of the bootstrap parameter estimates.

## Equal-Variance T-Test and Aspin-Welch Unequal-Variance Sections

These two sections present the t-test results for the equal variance and unequal variance cases, respectively. The definitions are essentially identical for each case, so they are combined.

| Alternative Hypothesis | TValue | Prob Level | Reject H0 at .050 | Power (Alpha=.05) | Power (Alpha=.01) |
|---|---|---|---|---|---|
| Difference <> 0 | -.1588 | .875032 | No | .052693 | .010837 |
| Randomization Test | | .879000 | | | |
| Difference < 0 | -.1588 | .437516 | No | .068110 | .014804 |
| Difference > 0 | -.1588 | .562484 | No | .035954 | .006616 |

Difference: (YldA)-(YldB)

**Aspin-Welch Unequal-Variance Test Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 | Power (Alpha=.05) | Power (Alpha=.01) |
|---|---|---|---|---|---|
| Difference <> 0 | -.1514 | .881278 | No | .052376 | .010723 |
| Difference < 0 | -.1514 | .440639 | No | .066968 | .014437 |
| Difference > 0 | -.1514 | .559361 | No | .036649 | .006802 |

Difference: (YldA)-(YldB)

### Alternative Hypothesis

This value identifies the test direction of the test reported in this row. Strict procedure requires you to select the null and alternative hypothesis prior to your analysis.

X-Y<>0. This is the two-tail test case. The null and alternative hypotheses are

$H_o$: $\mu_X$-$\mu_Y$=0, $H_a$: $\mu_X$-$\mu_Y \neq 0$.

X-Y<0. This is the left-tail test case. The null and alternative hypotheses are

$H_o$: $\mu_X$-$\mu_Y$=0, $H_a$: $\mu_X$-$\mu_Y$<0.

X-Y>0. This is the right-tail test case. The null and alternative hypotheses are

$H_o$: $\mu_X$-$\mu_Y$=0, $H_a$: $\mu_X$-$\mu_Y$>0.

### T-Value

This is the t-test test statistic computed from your data. The formulas for the two possible variance assumptions are identical in form. The only difference between them is that the appropriate denominator must be selected, depending on whether the variances are equal.

$$t = \frac{(\bar{x}_X - \bar{x}_Y) - (\mu_X - \mu_Y)}{s_{\bar{X}-\bar{Y}}}$$

### Prob Level

This is the p-value (significance level) for the statistical test. The p-value is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the p-value is less than $\alpha$, say 0.05, the null hypothesis is rejected. If the p-value is greater than $\alpha$, the null hypothesis is accepted.

### Reject H0 at .050

This is the conclusion reached (accept or reject) about the null hypothesis, H0. If you reject the null hypothesis, your study is said to be ***significant***. Otherwise, it is ***not significant***.

### Power (Alpha=0.05) and Power (Alpha=0.01)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error ($\beta$). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual difference between the two population means.

The power value calculated here assumes that the population standard deviation is equal to the sample standard deviation and that the difference between the population means is exactly equal to the difference between the sample means. Of course, this cannot be true, but it allows us to calculate the power at these values.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

Usually you would consider the power of the test when you failed to reject the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If you accept the null hypothesis with high power, there is not much left to do. At least you know that the two means are not different. However, if you accept the null hypothesis with low power, you can take one or more of the following actions:

1. Increase your alpha level. Perhaps you should be testing at alpha = 0.05 instead of alpha =.01. Increasing the alpha level will increase the power.
2. Increasing your sample size will increase the power of your test if you have low power. If you have high power, an increase in sample size will have little effect.
3. Decrease the magnitude of the variance. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

## Tests of Assumptions Section

This section presents the results of tests validating the normality and equal variance assumptions. Note that the t-test assumes that each group is normally distributed, so the normality tests are conducted on each group separately. Other assumptions concerning independence and random sampling are not tested here. You must justify those assumptions by considering your experiment procedure.

When using this report, all you need to do is scan down the column labeled Decision(5%). If none of the tests are rejected, you can feel confident that the assumptions are met. (Of course, the power of these tests is also influenced by your sample size. If you have a small sample size, say less than 25 per group, the power of these normality tests will be questionable and you will have to rely on other means to justify your assumptions.)

Two aspects of normality are tested for, skewness and kurtosis. If the normality of a batch of data fails because of skewness, it might be possible to use the square root or logarithmic transformation to normalize your data. If only one of the variables is normally distributed, look at the normal probability plot or box plot for the one variable that is not normally distributed to see if an outlier or two may have caused the nonnormality.

There are several schools of thought on whether a preliminary test for variance equality is proper before using the t-test. Various simulation studies that used the preliminary variance test have shown it to be very inadequate for a preliminary test. Our suggestion is to use the equal variance t-test when the sample sizes are equal or approximately equal and use the unequal variance t-test when the sample sizes are unequal. When the sample sizes are different, the most serious situation is when the smaller sample is associated with the larger variance. The other option is to use a different test for equality of variances. Conover and others (1981) did extensive simulation involving different distributions, sample sizes, means, and variances; and they found that the modified-Levene test is one of the most robust and powerful tests for equality of variance. Thus, if a preliminary test is to be preferred, use the modified-Levene test. Otherwise, do not do any preliminary test, and choose the t-test based on whether the sample sizes are equal.

In the case of nonnormality, two nonparametric tests were suggested. The basic assumptions of independent samples, continuous random variables, and a measurement scale of at least ordinal scale hold for both tests. The Mann-Whitney U or Wilcoxon Rank-Sum test has the additional assumption that the distributions for the two variables are identical (although not necessary normal) in form and shape (i.e., same variance) but differ only in location (i.e., in medians). On the other hand, the Kolmogorov-Smirnov is a general test for differences between two groups. As a general test, it is somewhat sensitive to all kinds of differences between groups or populations and yet not particularly sensitive to any specific type of difference. The Kolmogorov-Smirnov test is a good choice when there are a lot of ties in your data that tends to invalidate the Wilcoxon Rank-Sum test.

Finally, you should back up the results of these numerical tests by considering the box plots, histograms, and probability plots of the two groups. As explained below, they let you visually

determine if the assumptions of normality (probability plots) and equal variance (box plots) are justified.

| Assumption | Value | Probability | Decision(5%) |
|---|---|---|---|
| Skewness Normality (YldA) | 0.2691 | 0.787854 | Cannot reject normality |
| Kurtosis Normality (YldA) | 0.3081 | 0.758028 | Cannot reject normality |
| Omnibus Normality (YldA) | 0.1673 | 0.919743 | Cannot reject normality |
| Skewness Normality (YldB) | 0.4587 | 0.646444 | Cannot reject normality |
| Kurtosis Normality (YldB) | 0.1291 | 0.897258 | Cannot reject normality |
| Omnibus Normality (YldB) | 0.2271 | 0.892665 | Cannot reject normality |
| Variance Ratio Equal-Variance Test | 2.6020 | 0.083146 | Cannot reject equal variances |
| Modified Levene Equal-Variance Test | 1.9940 | 0.169347 | Cannot reject equal variances |

### Skewness Normality

This is a skewness test reported by D'Agostino (1990). Remember that skewness is lack of symmetry. One characteristic of the normal distribution is that it has no skewness. Hence, one type of nonnormality is skewness.

The Value is the test statistic for skewness, while Probability is the p-value for a two-tail test for normality. If this p-value is less than a chosen level of significance, usually 0.05, the data are not normally distributed according to this test. If the p-value is greater than the chosen level of significance, the data are assumed to be normally distributed. Under Decision (5%), the conclusion about normality is given.

### Kurtosis Normality

Kurtosis measures the heaviness of the tails of the distribution. D'Agostino (1990) reported a second normality test that tests kurtosis. The Value column gives the test statistic for kurtosis, while Probability is the p-value for a two-tail test for normality. If this p-value is less than a chosen level of significance, 0.05, the data are not normally distributed according to this test. If the p-value is greater than the chosen level of significance, the data are assumed normal. Under Decision (5%), the conclusion of the test is given. If the data are not normally distributed, the conclusion is rejection. If the data are normally distributed, the conclusion is acceptance.

### Omnibus Normality

This third normality test, also developed by D'Agostino (1990), combines the skewness and kurtosis tests into a single measure. The definitions for Value, Probability, and Decision (5%) are the same as for the previous two normality tests. This normality test is considered to be the best of the three.

### Variance Ratio Equal-Variance Test

This equal variance test is the ratio of two sample variances. This variance ratio is distributed as an F distribution with $n_x$-1 degrees of freedom for the numerator sample variance and $n_y$-1 degrees of freedom for the denominator sample variance.

$$F = \frac{s_X^2}{s_Y^2}$$

This variance ratio is shown under Value. Be careful! This test requires that the two samples are drawn from normal populations. If the two samples are not normally distributed, do not use this test as a preliminary test for equality of variances. It would be better to check for equality of variance with the modified Levene test or to use some graphic tool, such as the box plot.

The p-value (Probability) is compared to the level of significance. If it is less than the level of significance, there is a difference in variances and the Decision is rejection of the null hypothesis

of equal variances. If the p-value is greater than the level of significance, there is acceptance of equal variances.

### Modified Levene Equal-Variance Test

The modified Levene test has been found to be one of the best tests for equality of variances. Levene's procedure is outlined as follows. First, redefine the variables for each treatment or sample by taking the absolute value of the difference from the sample median. For one sample, this redefinition would be

$$z_{1j} = \left| x_j - Med_x \right|$$

And for the other,

$$z_{2j} = \left| y_j - Med_y \right|$$

Next, run a two-group one-way analysis of variance on this redefined variable. The F-value for this one-way analysis of variance is shown under Value and its corresponding p-value under Probability.

The p-value (Probability) is compared to the level of significance. If it is less than the level of significance, there is a difference in variances and the Decision is rejection of the null hypothesis of equal variances. Otherwise, there is acceptance of equal variances.

## Mann-Whitney U or Wilcoxon Rank-Sum Test

This test is the nonparametric substitute for the equal-variance t-test when the assumption of normality is not valid. When in doubt about normality, play it safe and use this test. The assumptions for this test were given in the Assumptions Section at the beginning of this chapter. Two key assumptions that we remind you of are that the distributions are at least ordinal in nature and that they are identical, except for location. This means that ties (repeated values) are not acceptable. When ties are present in your data, you can use the approximation provided, but know that the exact results no longer hold.

This particular test is based on ranks and has good properties (asymptotic relative efficiency) for symmetric distributions. There are exact procedures for this test given small samples with no ties, and there are large sample approximations.

| Variable | Mann Whitney U | W Sum Ranks | Mean of W | Std Dev of W |
|---|---|---|---|---|
| YldA | 101.5 | 192.5 | 195 | 22.79508 |
| YldB | 106.5 | 242.5 | 240 | 22.79508 |

Number Sets of Ties = 3, Multiplicity Factor = 18

| | Exact Probability | | Approximation Without Correction | | | Approximation With Correction | | |
|---|---|---|---|---|---|---|---|---|
| Alternative Hypothesis | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 | Z-Value | Prob Level | Reject H0 at .050 |
| Diff<>0 | | | -.1097 | .912671 | No | .0877 | .930086 | No |
| Diff<0 | | | -.1097 | .456336 | No | -.0877 | .465043 | No |
| Diff>0 | | | -.1097 | .543664 | No | -.1316 | .552351 | No |

### Variable

This is the name for each sample, group, or treatment.

### Mann Whitney U

The Mann-Whitney test statistic, U, is defined as the total number of times a Y precedes an X in the configuration of combined samples Gibbons (1985). It is directly related to the sum of ranks. This is why this test is sometimes called the Mann-Whitney U test and other times called the Wilcoxon Rank Sum test. The Mann-Whitney U test calculates Ux and Uy. The formula for Ux is as follows (the formula for Uy is obtained by replacing "x" by "y" in this formula):

$$Ux = Wx - \frac{n_x(n_x+1)}{2}$$

### W Sum Ranks

Given that the two samples (X and Y) are combined into one and the observations are ranked in ascending order, W is the sum of the ranks for the group or treatment, X or Y. Note that tied values are resolved by using the average rank of the tied values.

$$\text{Wx} = \Sigma(\text{ranks}_x)$$

### Mean of W

This is the mean of the distribution of W, formulated as follows:

$$\overline{W_x} = \frac{n_x(n_x+n_y+1)}{2}$$

and

$$\overline{W_y} = \frac{n_y(n_x+n_y+1)}{2}$$

### Std Dev of W

This is the standard deviation of the W corrected for ties. If there are no ties, this standard deviation formula simplifies since the second term under the radical is zero.

$$\sigma_W = \sqrt{\frac{n_x n_y(n_x+n_y+1)}{12} - \frac{n_x n_y \sum_{i=1}(t_i^3 - t_i)}{12(n_x+n_y)(n_x+n_y-1)}}$$

where $t_1$ is the number of observations tied at value one, $t_2$ is the number of observations tied at some value two, and so forth. Generally, this correction for ties in the standard deviation makes little difference unless there are a lot of ties.

### Number Sets of Ties

This gives the number of sets of tied values. If there are no ties, this number is zero. A set of ties is two or more observations with the same value. This is used in adjusting the standard deviation for the W.

### Multiplicity factor

This is the tie portion of the standard deviation of W, given by

$$\sum_{i=1}(t_i^3 - t_i)$$

### Alternative Hypothesis

For the Wilcoxon rank-sum test, the null and alternative hypotheses relate to the equality or non-equality of two medians. If a difference other than zero is desired between the medians of the null

hypothesis (such as $H_o$: $median_x$=$median_y$ + *d* where *d* is some specified number), simply add the number *d* to each Y, and run the test on the original X's and the newly adjusted Y's.

This value identifies the test direction of the test reported in this row. Strict statistical procedure requires you to select the null and alternative hypothesis prior to your analysis.

Diff<>0. This is the two-tail test case. The null and alternative hypotheses are

$H_o$: $Median_X$=$Median_Y$ , $H_a$: $Median_X \neq Median_Y$.

Diff<0. This is the left-tail test case. The null and alternative hypotheses are

$H_o$: $Median_X = Median_Y$ , $H_a$: $Median_X < Median_Y$

Diff>0. This is the right-tail test case. The null and alternative hypotheses are

$H_o$: $Median_X = Median_Y$, $H_a$: $Median_X > Median_Y$

### Exact Probability: Prob Level

This is an exact p-value for this statistical test based on the distribution of W. This p-value assumes no ties (if ties are detected, this value is left blank). The p-value is the probability that the test statistic will take a value at least as extreme as the actually observed value, assuming that the null hypothesis is true. The exact probability value is available for sample sizes up to 38.

### Exact Probability: Reject H0 at .050

This is the conclusion reached about the null hypothesis.

### Approximation without correction: Z value

A normal approximation method can be used for the distribution of the sum of ranks which corrects for ties but does not have the correction factor for continuity. The z value is:

$$z = \frac{Wn - \mu_{Wn}}{\sigma_{Wn}}$$

where W is the sum of ranks for the smaller sample size and $\mu_W$ is the mean of W. The z value, the p-value, and the decision at specified alpha level are provided.

### Approximation with correction: Z value

This is a normal approximation that corrects for ties and has the correction factor for continuity. The *z* value is:

$$z = \frac{Wn - \mu_{Wn} + 0.5}{\sigma_{Wn}}$$

where W is the sum of ranks for the smaller sample size and $\mu_W$ is the mean of W.

If a normal approximation procedure is used, this one is the most accurate.

### Prob Level

This is the p-value for the Wilcoxon rank-sum test. Exact values are given for sample sizes under 40. The normal approximation approach is reported for sample sizes over 40. The p-value is the probability that the test statistic will take a value at least as extreme as the actually observed value, assuming that the null hypothesis of equality of medians is true. If the p-value is less than $\alpha$, say 0.05, the null hypothesis is rejected. If the p-value is greater than $\alpha$, the null hypothesis is accepted.

### Reject H0 at .050

This is the conclusion about H0 that is reached.

## Kolmogorov-Smirnov Test

This is a two-sample test for differences between two samples or distributions. If a statistical difference is found between the distributions of X and Y, the test provides no insight as to what caused the difference. For example, the difference could be due to differences in location (mean), variation (standard deviation), presence of outliers, type of skewness, type of kurtosis, number of modes, and so on.

The assumptions for this nonparametric test are: (1) there are two independent random samples; (2) the two population distributions are continuous; and (3) the data are at least ordinal in scale. This test is frequently preferred over the Wilcoxon sum-rank test when there are a lot of ties. The test statistic is the maximum distance between the empirical distribution functions (EDF) of the two samples.

| Alternative Hypothesis | Dmn Criterion Value | Reject Ho if Greater Than | Test Alpha Level | Reject H0 at .050 | Prob Level |
|---|---|---|---|---|---|
| D(1)<>D(2) | 0.322115 | 0.4768 | .050 | No | .3468 |
| D(1)<D(2) | 0.322115 | 0.4768 | .025 | No | |
| D(1)>D(2) | 0.177885 | 0.4768 | .025 | No | |

### Alternative Hypothesis

The null and alternative hypotheses relate to the equality of the two distribution functions (noted as F(X) or F(Y)). This value identifies the test direction of the test reported in this row. Strict procedure requires you to select the null and alternative hypotheses prior to your analysis.

D(1)<>D(2). This is the two-tail test case. The null and alternative hypotheses are

$$H_o: F(X) = F(Y), H_a: F(X) \neq F(Y)$$

D(1)<D(2). This is the left-tail test case. The null and alternative hypotheses are

$$H_o: F(X) = F(Y), H_a: F(X) < F(Y)$$

D(1)>D(2). This is the right-tail test case. The null and alternative hypotheses are

$$H_o: F(X) = F(Y), H_a: F(X) > F(Y)$$

### Dmn-criterion value

This is the maximum difference between the two empirical distribution functions. It is the Kolmogorov-Smirnov test statistic.

### Reject $H_o$ if Greater Than

This number is the decision criterion for the Kolmogorov-Smirnov test based on $n_X$ and $n_Y$. If the test statistic *Dmn* is greater than this decision limit, there is a statistically significant difference between the two samples. However, we do not know what aspect of the two samples is different.

### Test Alpha Level

This is the level of significance, $\alpha$, for this test.

### Reject H0 at .050

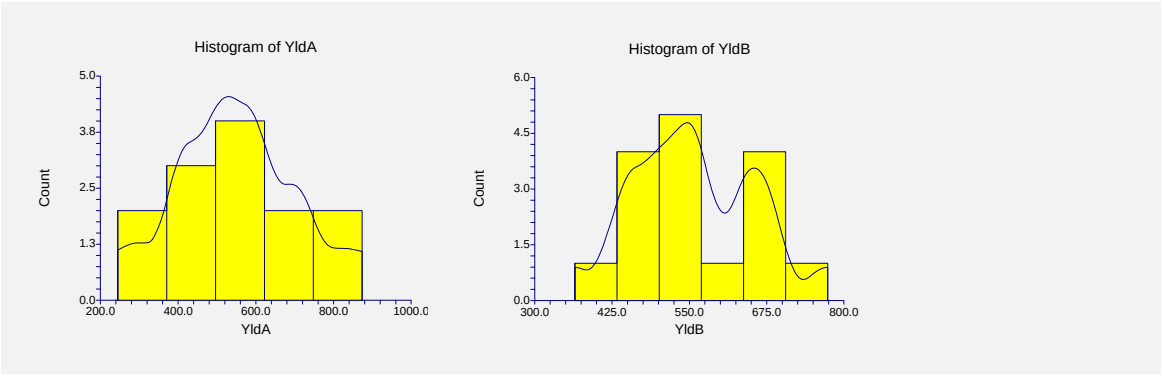
If the level of significance is α, a No means that the test statistic was less than the decision criterion and that there is no statistical difference between the two samples. A Yes means that there is a statistical difference between the two groups.

### Prob Level

This is the p-value for a two-tail test. If the level of significance, α, is larger than this p-value, reject H0.

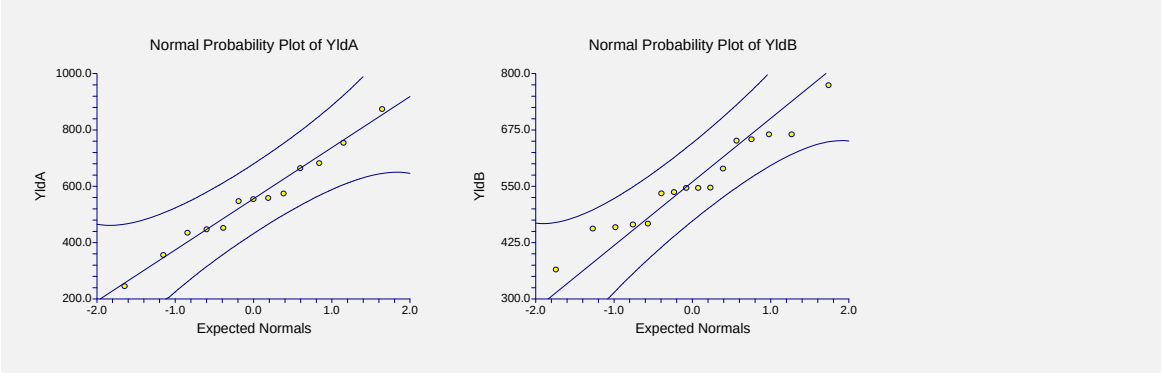
## Histogram and Density Trace

The histogram with the density trace overlay (the wavy line) lets you study the distributional features of the two samples to determine if (and which) two-sample tests are appropriate. Note that histograms require larger samples than probability plots. Since the shape of the histogram is influenced by the number of classes or bins and the width of the bins, the best choice is to trust the density trace, which is a smoothed histogram. A complete discussion of histograms is given in the chapter on this topic.



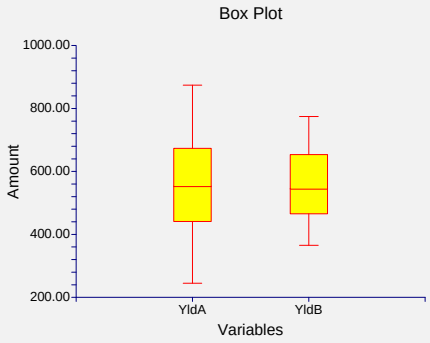

## Normal Probability Plots

This is a normal probability plot of the actual data values for each sample. There would be two normal probability plots. If any of the data values fall outside the confidence bands, the data are not normal for that group. The goodness-of-fit tests mentioned earlier, especially the omnibus test, should confirm this fact statistically. If only one observation falls outside the confidence bands and the remaining data values hug the straight line, there is an outlier.



## Box Plots

Box plots are useful for assessing symmetry, presence of outliers, general equality of location, and equality of variation.




# Two-Sample T-Test Checklist

This checklist, prepared by a professional statistician, is a flowchart of the steps you should complete to conduct a valid two-sample t-test (or one of its nonparametric counterparts). You should complete these tasks in order.

## Step 1 – Data Preparation

### Introduction

This step involves scanning your data for anomalies, keypunch errors, typos, and so on. You would be surprised how often we hear of people completing an analysis, only to find that they had mistakenly selected the wrong database.

### Sample Size

The sample size (number of nonmissing rows) has a lot of ramifications. The two-sample t-test was developed under the assumption that the sample sizes of each group would be equal. In practice, this seldom happens, but the closer you can get to equal sample sizes the better.

With regard to the combined sample size, the t-test may be performed on very small samples, say 4 or 5 observations per group. However, in order to test assumptions and obtain reliable estimates of variation, you should attempt to obtain at least 30 individuals per group.

It is possible to have a sample size that is too large. When your sample size is quite large, you are almost guaranteed to find statistical significance. However, the question that then arises is whether the magnitude of the difference is of practical importance.

### Missing Values

The number and pattern of missing values are always issues to consider. Usually, we assume that missing values occur at random throughout your data. If this is not true, your results will be biased since a particular segment of the population is underrepresented. If you have a lot of

missing values, some researchers recommend comparing other variables with respect to missing versus nonmissing. If you find large differences in other variables, you should begin to worry about whether the missing values are cause for a systematic bias in your results.

### Type of Data

The mathematical basis of the t-test assumes that the data are continuous. Because of the rounding that occurs when data are recorded, all data are technically discrete. The validity of assuming the continuity of the data then comes down to determining when we have too much rounding. For example, most statisticians would not worry about human-age data that was rounded to the nearest year. However, if these data were rounded to the nearest ten years or further to only three groups (young, adolescent, and adult), most statisticians question the validity of the probability statements. Some studies have shown that the t-test is reasonably accurate when the data has only five possible values (most would call this discrete data). If your data contains less than five unique values, any probability statements made are tenuous.

### Outliers

Generally, outliers cause distortion in most popular statistical tests. You must scan your data for outliers (the box plot is an excellent tool for doing this). If you have outliers, you have to decide if they are one-time occurrences or if they would occur in another sample. If they are one-time occurrences, you can remove them and proceed. If you know they represent a certain segment of the population, you have to decide between biasing your results (by removing them) or using a nonparametric test that can deal with them. Most would choose the nonparametric test.

## Step 2 – Setup and Run the T-Test Panel

### Introduction

Now comes the fun part: running the program. ***NCSS*** is designed to be simple to operate, but it can still seem complicated. When you go to run a procedure such as this for the first time, take a few minutes to read through the chapter again and familiarize yourself with the issues involved.

### Enter Variables

The ***NCSS*** procedures are set with ready-to-run defaults. About all you have to do is select the appropriate variables.

### Select All Plots

As a rule, you should select all diagnostic plots (box plots, histograms, etc.) even though they may take a few extra seconds to generate. They add a great deal to your analysis of the data.

### Specify Alpha

Most beginners in statistics forget this important step and let the alpha value default to the standard 0.05. You should make a conscious decision as to what value of alpha is appropriate for your study. The 0.05 default came about when people had to rely on printed probability tables and there were only two values available: 0.05 or 0.01. Now you can set the value to whatever is appropriate.

# Step 3 – Check Assumptions

## Introduction

Once the program output is displayed, you will be tempted to go directly to the probability of the t-test, determine if you have a significant result, and proceed to something else. However, it is very important that you proceed through the output in an orderly fashion. The first task is to determine which assumptions are met by your data.

Sometimes, when the data are nonnormal for both samples, a data transformation (like square roots or logs) might normalize the data. Frequently, when only one sample is normal and the other is not, this transformation, or re-expression, approach works well.

It is not unusual in practice to find a variety of tests being run on the same basic null hypothesis. That is, the researcher who fails to reject the null hypothesis with the first test will sometimes try several others and stop when the hoped-for significance is obtained. For instance, a statistician might run the equal-variance t-test on the original two samples, the equal-variance t-test on the logarithmically transformed data, the Wilcoxon rank-sum test, and the Kolmogorov-Smirnov test. An article by Gans ("The Search for Significance:  Different Tests on the Same Data," *The Journal of Statistical Computation and Simulation*," 1984, pp. 1-21) suggests that there is no harm on the true significance level if no more than two tests are run. This is not a bad option in the case of questionable outliers. However, as a rule of thumb, it seems more honest to investigate whether the data are normal. The conclusion from that investigation should direct one to the right test.

## Random Sample

The validity of this assumption depends upon the method used to select the sample. If the method used assures that each individual in the population of interest has an equal probability of being selected for this sample, you have a random sample. Unfortunately, you cannot tell if a sample is random by looking at it or statistics from it.

## Sample Independence

The two samples must be independent. For example, if you randomly divide a group of individuals into two groups, you have met this requirement. However, if your population consists of cars and you assign the left tire to one group and the right tire to the other, you do not have independence. Here again, you cannot tell if the samples are independent by looking at them. You must consider the sampling methodology.

## Check Descriptive Statistics

You should check the Individual-Group Statistics Section first to determine if the Count and the Mean are reasonable. If you have selected the wrong variable, these values will alert you.

## Normality

To validate this assumption, you would first look at the plots. Outliers will show up on the box plots and the probability plots. Skewness, kurtosis, more than one mode, and a host of other problems will be obvious from the density trace on the histogram. No data will be perfectly normal. After considering the plots, look at the Tests of Assumptions Section to get numerical confirmation of what you see in the plots. Remember that the power of these normality tests is

directly related to the sample size, so when the normality assumption is accepted, double-check that your sample is large enough to give conclusive results.

### Equal Variance

The equal variance assumption is important in determining which statistical test to use. Check the box plots for boxes with about the same widths. Confirm your conclusion by looking at the Equal-Variance Test (Modified Levene) line. Note that, strictly speaking, these equal variance tests require the assumption of normality. If your data are not normal, you should use the modified Levene test. It works in many nonnormal situations.

Some researchers recommend against using a preliminary test on variances (which research and simulations do not strongly support). If you decide against these preliminary tests, base your choice of a test procedure on the sample sizes. If the two sample sizes are approximately equal, use the equal-variance t-test. If the ratio of the two sample sizes (larger sample size over the smaller sample size) is equal to or greater than 1.5, use the unequal-variance t-test. This is the recommendation of Ott (1984), page 144.

## Step 4 – Choose the Appropriate Statistical Test

### Introduction

After understanding how your data fit the assumptions of the various two-sample tests, you are ready to determine which statistical procedures will be valid. You should select one of the following four situations based on the status of the normality and equal variance assumptions.

### Normal Data with Equal Variances

Use the Equal Variance T-Test Section for hypothesis testing and the Equal Variance portion of the Confidence Limits Section for interval estimation.

### Normal Data with Unequal Variances

Use the Unequal Variance T-Test Section for hypothesis testing and the Unequal Variance portion of the Confidence Limits Section for interval estimation.

### Non-Normal Data with Equal Variances

Use the Mann-Whitney U or Wilcoxon Rank-Sum Test for hypothesis testing.

### Non-Normal Data with Unequal Variances

Use the Kolmogorov-Smirnov Test in this case or if your data have a lot of ties.

## Step 5 – Interpret Findings

### Introduction

You are now ready to conduct your two-sample test. Depending upon the nature of your study, you look at either of the following sections.

### Hypothesis Testing

First find the appropriate Alternative Hypothesis row. Usually, you will use the first (Var1-Var2<>0) row. This two-tailed test is the standard. If the probability level is less than your chosen alpha level, you reject the null hypothesis of equal means (or medians) and conclude that the means are different. Your next task is to look at the means themselves to determine if the size of the difference is of practical interest.

### Confidence Limits

The confidence limits of the difference let you put bounds on the size of the difference. If these limits are narrow and close to zero, you might determine that even though your results are statistically significant, the magnitude of their difference is not of practical interest.

## Step 6 – Record Your Results

Finally, as you finish a test, take a moment to jot down your impressions. Explain what you did, why you did it, what conclusions you reached, which outliers you deleted, areas for further investigation, and so on. Since this is a technical process, your short-term memory will not retain it for long. These notes will be worth their weight in gold when you come back to this printout a few days later!

# Example of Two-Sample T-Test Steps

This example illustrates the interpretation of two-sample tests. Of course, no example is infallible, but the intention is to highlight a number of the issues that you must consider in choosing the right two-sample test for your data as you proceed through the Two-Sample Checklist.

Two friends, who are also neighbors, love pizza, and they each usually order their pizzas from different places. Friend A orders from pizza company 1, while friend B orders from pizza company 2. The two friends got in an argument about which pizza company delivers the fastest or whether there was a difference at all in delivery times. Friend A took a random sample of 10 delivery times from pizza place 1 over the next six months. Friend B took a random sample of 8 delivery times over the same time frame. The pizza orders were not necessarily taken on the same day, but the orders were usually placed in the evening hours from 6 to 9 p.m. The data are shown below.

| Pizza1 | Pizza2 |
|---|---|
| 21 | 15 |
| 20 | 17 |
| 25 | 17 |
| 20 | 19 |
| 23 | 22 |
| 20 | 12 |
| 13 | 16 |
| 18 | 21 |
| 25 | |
| 24 | |

## Step 1 – Data Preparation

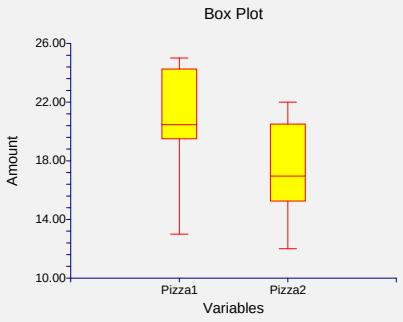
The sample sizes here are not as large as we would like, but they are typical. There are no missing values, and the data are continuous (although the times are rounded to the closest minute). There is no way to assess outliers until Step 3.

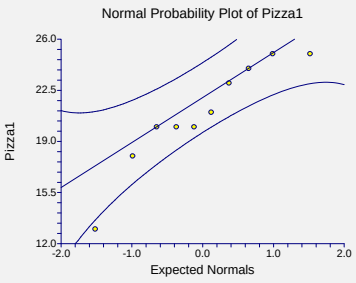
## Step 2 – Setup and Run the T-Test Panel

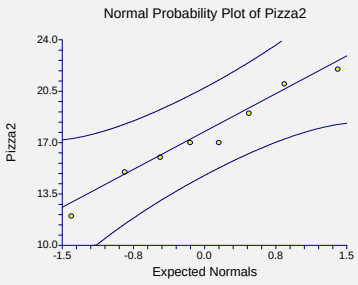
The selection and running of the Two-Sample T-Test from the Analysis menu on the two response variables, Pizza1 and Pizza2, would produce the reports that follow. The alpha value has been set at 0.05.

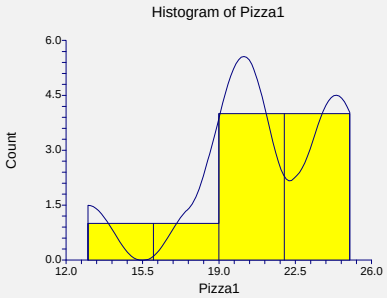
## Step 3 – Check Assumptions

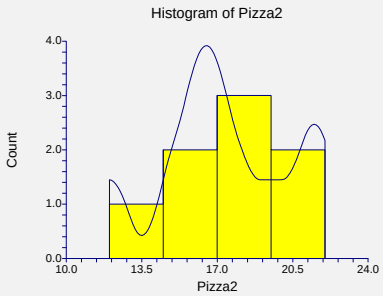
We first check for normality with the graphic perspectives: box plots, normal probability plots, histograms, and density traces.



The tails of the box plot for Pizza1 show left skewness, and the median is not in the middle of the box itself (i.e., it is also pulled left). While the tails for Pizza2 are symmetrical, the median is also pulled left toward the short delivery times. Remember that these samples are small, and interpretation of box plots for small samples must be flexible. The interpretation from the box plots is that both groups show some non-normality.

The normal probability plots in give a similar picture. Since all of the data values for Pizza2 lie within the 95% confidence bands, delivery times seem to be normal. On the other hand, the normal probability plot for Pizza1 shows a possible outlier among the short delivery times since the observation of 13 minutes is outside the confidence bands. If it were not for this one observation, the normal probability plot for Pizza1 would be normal.

The histogram does not usually give an accurate graphic perception of normality for small samples, although the super-imposed density trace helps a lot. Examination of the histogram for Pizza1 shows that there is at least one observation that contributes to the left skewness, and the histogram for Pizza1 does not look normal. However, the histogram for Pizza2 reveals a reasonably normal distribution.

At this point of the graphic analysis of the normality assumption, you would likely say the Pizza2 delivery times are normal while Pizza1 delivery times are not. However, since these samples are small, be sure to evaluate the numerical confirmation of normality by the skewness, kurtosis, and omnibus normality tests for each pizza firm using the Tests of Assumptions Section.

**Test of Assumptions Section**

| Assumption | Value | Probability | Decision(5%) |
|---|---|---|---|
| Skewness Normality (pizza1) | -1.4244 | 0.154336 | Cannot reject normality |
| Kurtosis Normality (pizza1) | 1.0525 | 0.292585 | Cannot reject normality |
| Omnibus Normality (pizza1) | 3.1366 | 0.208404 | Cannot reject normality |
| Skewness Normality (pizza2) | -0.1541 | 0.877509 | Cannot reject normality |
| Kurtosis Normality) (pizza2) | -0.0171 | 0.986393 | Cannot reject normality |
| Omnibus Normality (pizza2) | 0.0240 | 0.988049 | Cannot reject normality |
| Variance Ratio Equal-Variance Test | 1.2729 | 0.767280 | Cannot reject equal variances |
| Modified Levene Equal-Variance Test | 0.0945 | 0.762487 | Cannot reject equal variances |

When evaluating normality, focus your attention on the probability (p-value) and the decision for the given alpha of 0.05. In this case, the decision is acceptance of the hypothesis that the data for Pizza1 is normally distributed by all three normality tests. The lowest probability is 0.1543 for the skewness test, and this is greater than 0.05, the set alpha value. This same amount of skewness for a larger sample size would have rejected the normality assumption. However, for our example, it seems reasonable to assume that both Pizza1 and Pizza2 are normally distributed. We would strongly recommend that the one outlying value in Pizza1 be double-checked for validity.

We next check for equal variance. Both variance tests (variance-ratio and modified-Levene) indicate acceptance of the hypothesis of equal variances as a shown by the probability greater than 0.05 and the “cannot reject” under the decision conclusion. This equality of variances is portrayed by the box plots.

If you do not consider the preliminary test on variances appropriate, use the sample size criterion. If the sample sizes are roughly equal (no more than a 1.5 ratio of the larger sample size divided by the smaller sample size), use the equal-variance t-test. In this case, this sample size ratio is 10/8 or 1.25. Thus, go ahead with the equal variance t-test. If you are in doubt, run both tests and compare answers.

## Step 4 – Choose the Appropriate Statistical Test

In this example, the conclusions from the assumption checking have been that both samples are normally distributed and that the variances are equal or that the sample sizes are roughly equal. In light of these findings, the appropriate test is the equal-variance t-test, sometimes called the pooled t-test.

## Step 5 – Interpret Findings

In order to understand the following discussion, you should run the two-sample t-test on the above data and look at the statistical reports.

The mean delivery times are 20.9 and 17.4 minutes. Note that the standard deviations are about equal at 3.665 and 3.249 minutes for Pizza1 and Pizza2, respectively.

We are interested in the difference between the means. Under the Confidence Limits Section and the Equal Variance Case, the 95% confidence limits for the difference ranges from 0.016557 to 7.033442 minutes. Since zero is not in this interval, there is a statistically significant difference between the two means.

The formal two-tail hypothesis test for this example is shown under the Equal-Variance T-Test section. The p-value or probability of accepting $H_o$ is 0.049, which is less than the chosen alpha level at 0.05, resulting in the rejection of $H_o$. That is, there is a difference between the two pizza delivery times. The power of this two-tail t-test at 0.05 level of significance is 0.5166. The higher the power (i.e., closer to 1), the better the statistical test is able to detect that the alternative hypothesis is true. The power is not great here (many would find it bearable), and it could have been greatly improved by slightly larger sample sizes.

If we had been interested in checking for the average Pizza1 delivery times being greater than that of Pizza2, we would have looked at the right-tail test in the equal-variance t-test section. The decision here is definitely a rejection since the p-value or the probability of accepting $H_o$ is significantly less than 0.05 (i.e., 0.0245). The power of this one-tail test is much better at 0.653.

This would usually finish the interpretation of this example. However, if you were having second thoughts about the normality for Pizza1 delivery times, you might check the nonparametric equivalent of the equal-variance t-test--the Mann-Whitney U Test--to see if you obtain a similar conclusion. The approximate p-value for the two-tail test is 0.044. This p-value is close to that which we had under the equal-variance t-test. Note that we still reject the null hypothesis. The right-tail test yields a p-value of 0.022, which is almost identical to the equal-variance t-test p-value for this right tail test.

Whenever the data are normal, use the appropriate t-test because the power is always better. If in doubt, cross check your t-test with the appropriate nonparametric test.

This concludes the analysis of this example.

Chapter 207

# T-Test – Two-Sample (From Means and SD's)

## Introduction

This program computes the two-sample t-test directly from the mean, standard deviation, and sample size. Confidence intervals for the means, difference, and standard deviation are computed. Hypothesis tests include the results for both one and two sided tests as well as equivalence tests.

## Technical Details

The formulas used by this procedure are the same as those presented in the Two-Sample T-Test. We refer you to that chapter for details. In this section, technical details of new output not documented previously are added.

### Equivalence Tests

An equivalence test is designed to show that one (new) treatment is similar to, but not necessarily better than, another (standard) treatment. To accomplish this, the roles of the null and alternative hypotheses are reversed. The hypotheses for testing equivalence of two means are (assuming that $\delta_L < 0$ and $\delta_U > 0$)

$$H_0: \mu_1 - \mu_2 \le \delta_L \quad \text{or} \quad \mu_1 - \mu_2 \ge \delta_U \quad \text{versus} \quad H_1: \delta_L < \mu_1 - \mu_2 < \delta_u$$

The alternative hypothesis states that the true difference is in some small, clinically acceptable range. For example, we might be willing to conclude that the benefits of two drugs are equivalent if the difference in their mean response rates is between -0.1 and 0.1.

The conventional method of testing equivalence hypotheses is to perform two, one-sided tests (TOST) of hypotheses. The null hypothesis of non-equivalence is rejected in favor of the alternative hypothesis of equivalence if both one-sided tests are rejected. Unlike the common

two-sided tests, however, the type I error rate is set directly at the nominal level (usually 0.05)—it is not split in half. So, to perform the test, two, one-sided tests are conducted at the $\alpha$ significance level. If both are rejected, the alternative hypothesis is concluded at the $\alpha$ significance level. Note that the *p*-value of the test is the maximum of the *p*-values of the two tests.

The two, one-sided tests of hypotheses for the difference are

$$H_{01}: \mu_1 - \mu_2 \le \delta_L \quad \text{versus} \quad H_{11}: \mu_1 - \mu_2 > \delta_L$$
$$H_{02}: \mu_1 - \mu_2 \ge \delta_U \quad \text{versus} \quad H_{12}: \mu_1 - \mu_2 < \delta_U$$

## Confidence Intervals for the Standard Deviation

Using the common notation for sample statistics (see, for example, ZAR (1984) page 115), a $100(1-\alpha)\%$ confidence interval for the standard deviation is given by

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}} \le \sigma \le \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}}$$

Note that this interval relies heavily on the assumption that the underlying data distribution is normal. If the data distribution is not normal, you should not use these results.

## Confidence Intervals for the Standard Deviation Ratio

Using the common notation for sample statistics (see, for example, ZAR (1984) page 125), a $100(1-\alpha)\%$ confidence interval for the ratio of two standard deviations is given by

$$\frac{s_1}{s_2\sqrt{F_{1-\alpha/2,n_1-1,n_2-1}}} \le \frac{\sigma_1}{\sigma_2} \le \frac{s_1\sqrt{F_{1-\alpha/2,n_2-1,n_1-1}}}{s_2}$$

Note that this interval relies heavily on the assumption that the underlying data distribution is normal. If the data distribution is not normal, you should not use these results.

# Data Structure

This procedure does not use data from the database. Instead, you enter the values directly into the panel.

# Procedure Options

This section describes the options available in this procedure.

## Data Tab

Enter the data values directly on this panel.

### Groups 1 and 2

#### N1 and N2 (Sample Size)

These boxes specify the sample sizes (number of subjects) in each group.

#### M1 and M2 (Means)

These boxes specify the sample means for each group. Note that you should enter as many digits as possible—do not round off the value if possible.

#### SD1 and SD2 (Standard Deviations)

These boxes specify the sample standard deviation for each group. Note that you should enter as many digits as possible—do not round off the value if possible.

Note that you must enter either the standard deviation (SD) or the standard error (SE), but not both.

#### SE1 and SE2 (Standard Errors of the Mean)

These boxes specify the sample standard errors for each group. Note that you should enter as many digits as possible—do not round off the value if possible.

Note that you must enter either the standard deviation (SD) or the standard error (SE), but not both.

### Options

#### H0: Diff. (M1-M2)

Enter the hypothesized value of the difference between mean $\delta_0$ under the null hypothesis. Usually, this value is zero. This option lets you specify a value other than zero, which is commonly used for non-inferiority tests.

#### Upper and Lower Equivalence Bounds

These options specify the upper and lower equivalence bounds for the test of mean equivalence. That is, these options specify $\delta_U$ and $\delta_L$. Usually, $\delta_L = -\delta_U$, but this is not required.

This value is sometimes called the *margin of equivalence*. It represents the largest difference that would still result in the conclusion of equivalence. For example, suppose that if the mean responses of two drugs are no more than 5 units apart, they are to be considered equivalent. Then, in this case, the margin of equivalence is 5.

If this value is left blank, the equivalence test is not displayed.

# Reports Tab

This panel contains options that control the format of the output.

## Report Options – Decimal Places

### Means, SD's ... Test Values

The number of digits displayed to the right of the decimal place.

## Report Options – Alpha Levels

### Confidence Limits

This option sets the alpha value for any confidence limits that are generated. The confidence coefficient of a confidence interval is equal to 1 - alpha. Thus, an alpha of 0.05 results in a confidence coefficient of 95%. Typical values are 0.01, 0.05, and 0.10.

### Hypothesis Test

This option sets the alpha value for any hypothesis tests that are generated. Typical values are 0.05, 0.01, and 0.10.

Note that alpha is the probability of rejecting the null hypothesis when it is true.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the

# Example 1 – Analyzing Summarized Data Using a T-Test

This section presents an example of using this panel to analyze a set of previously summarized data. A published report showed the following results: N1 = 15, Mean1 = 3.7122, SD1 = 1.9243, N2 = 13, Mean2 = 1.8934, and SD2 = 2.4531. Along with the other results, suppose you want to see an equivalence test in which the margin of equivalence is 0.3.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the T-Test – Two-Sample (From Means and SD's) window.

**1 Open the T-Tests - Two Sample (From Means and SD's) window.**

- On the menus, select **Analysis**, then **T-Tests**, then **T-Tests - Two Sample (From Means and SD's)**. The procedure will be displayed.

**2 Specify the data.**

- On the window, select the **Data tab**.
- In the **N1** box, enter **15**.
- In the **M1** box, enter **3.7122**.
- In the **SD1** box, enter **1.9243**.
- In the **N2** box, enter **13**.
- In the **M2** box, enter **1.8934**.
- In the **SD2** box, enter **2.4531**.
- In the **Upper Equiv. Bound** box, enter **0.3**.

**3 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports will be displayed in the Output window.

## Confidence Intervals of Means

| Sample | N | Mean | Standard Deviation | Standard Error | 95% Lower Conf. Limit of Mean | 95% Upper Conf. Limit of Mean |
|---|---|---|---|---|---|---|
| 1 | 15 | 3.7122 | 1.9243 | 0.4969 | 2.6466 | 4.7778 |
| 2 | 13 | 1.8934 | 2.4531 | 0.6804 | 0.4110 | 3.3758 |

This report documents the values that were input alone with the associated confidence intervals.

## Confidence Intervals of Difference

| Variance Assumption | DF | Mean Difference | Standard Deviation | Standard Error | 95% Lower Conf. Limit of Difference | 95% Upper Conf. Limit of Difference |
|---|---|---|---|---|---|---|
| Equal | 26.00 | 1.8188 | 2.1843 | 0.8277 | 0.1174 | 3.5202 |
| Unequal | 22.68 | 1.8188 | 3.1178 | 0.8425 | 0.0747 | 3.5629 |

This report provides confidence intervals for the difference between the means. The first row gives the equal-variance interval. The second row gives the interval corrected for unequal variances.

The interpretation of these confidence intervals is that when populations are repeatedly sampled and confidence intervals are calculated, 95% of those confidence intervals will include (cover) the true value of the difference.

## Hypothesis Tests

**Two-Sided, Two-Sample T-Test (H0: M1 = M2 versus H1: M1 <> M2)**

| Variance Assumption | DF | T-Value | Prob Level | Conclude H1 at 5.0% Significance? | Power (Alpha=0.05) | Power (Alpha=0.01) |
|---|---|---|---|---|---|---|
| Equal | 26.00 | 2.197 | 0.0371 | Yes | 0.5619 | 0.3025 |
| Unequal | 22.68 | 2.159 | 0.0417 | Yes | 0.5427 | 0.2834 |

**Lower, One-Sided, Two-Sample T-Test (H0: M1 = M2 versus H1: M1 < M2)**

| Variance Assumption | DF | T-Value | Prob Level | Conclude H1 at 5.0% Significance? | Power (Alpha=0.05) | Power (Alpha=0.01) |
|---|---|---|---|---|---|---|
| Equal | 26.00 | 2.197 | 0.9814 | No | 0.0001 | 0.0000 |
| Unequal | 22.68 | 2.159 | 0.9792 | No | 0.0001 | 0.0000 |

**Upper, One-Sided, Two-Sample T-Test (H0: M1 = M2 versus H1: M1 > M2)**

| Variance Assumption | DF | T-Value | Prob Level | Conclude H1 at 5.0% Significance? | Power (Alpha=0.05) | Power (Alpha=0.01) |
|---|---|---|---|---|---|---|
| Equal | 26.00 | 2.197 | 0.0186 | Yes | 0.6896 | 0.4040 |
| Unequal | 22.68 | 2.159 | 0.0208 | Yes | 0.6733 | 0.3838 |

These reports give the results of the t-test of whether the means are equal. The first line gives t-test based on the equal variance assumption. This is the standard t-test. The second line gives the t-test using an adjustment to compensate for unequal group variances.

The power values are computed assuming that the observed difference in the sample means coincides with the true difference in the population means.

### Variance Assumption

Two t-tests are conducted for each set of hypotheses. The equal-variance test is the classical t-test. The unequal-variance test is adjusted to compensate for unequal group variances.

### DF

This column specifies the degrees of freedom. Note that fractional degrees of freedom are usually obtained with the unequal-variance test.

### T-Value

This is value of the *t* statistic.

### Prob Level

This is the *p-value* (significance level) of the test. The p-value is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the p-value is less than 0.05, the null hypothesis is rejected. If the p-value is greater than 0.05, the null hypothesis is accepted.

### Conclude H1 at 5% Significance

This is the conclusion reached about the null hypothesis. When H0 is rejected, you conclude that H1 is true and the results are said to be *significant*. When H0 is not rejected, the results are said to be non-significant. Note that a non-significant result does not establish H0. If you wish to establish H0, you should use an equivalence test.

### Power (Alpha=0.05) and Power (Alpha=0.01)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error (β). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual difference between the two population means.

The power value calculated here assumes that the population standard deviation is equal to the sample standard deviation and that the difference between the population means is exactly equal to the difference between the sample means. Of course, this cannot be true, but it allows us to calculate the power at these values.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

Usually you would consider the power of the test when you failed to reject the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If you do not reject H0 and you have high power, there is not much left to do.

## Equivalence Tests of Means

**Equivalence Tests of Difference (H0: M1 - M2 < -0.3000 or M1 - M2 > 0.3000 versus H1: Equivalence)**
**Estimated Difference (M1 - M2) = 1.8188**

| Variance Assumption | Lower Test Statistic's Value | Lower Test Statistic's Prob | Upper Test Statistic's Value | Upper Test Statistic's Prob | Prob Level | Conclude H1 at 5.0% Significance? |
|---|---|---|---|---|---|---|
| Equal | 1.835 | 0.9610 | 2.560 | 0.0083 | 0.9610 | No |
| Unequal | 1.803 | 0.9576 | 2.515 | 0.0097 | 0.9576 | No |

This report gives the results the equivalence test. The equivalence test is designed to establish that, for practical purposes, the two means are equal.

### Lower Test Statistic's Value

The equivalence test is based on two, one-sided tests (TOST). This is the test statistic for the lower test.

### Lower Test Statistic's Probability

The equivalence test is based on two, one-sided tests (TOST). This is the significance level for the lower test.

### Upper Test Statistic's Value

The equivalence test is based on two, one-sided tests (TOST). This is the test statistic for the upper test.

### Upper Test Statistic's Probability

The equivalence test is based on two, one-sided tests (TOST). This is the significance level for the upper test.

### Prob Level

This is the significance level of the test. This value is the maximum of the lower and upper probabilities. If this value is less that 0.05, the null hypothesis of non-equivalence is rejected and equivalence is concluded.

### Conclude H1 at 5% Significance?

If this value is 'No', equivalence is not established. If this value is 'Yes', equivalence is established.

## Confidence Intervals of Standard Deviations

| Sample | DF | Mean Difference | Standard Deviation | Standard Error | 95% Lower Conf. Limit of SD | 95% Upper Conf. Limit of SD |
|---|---|---|---|---|---|---|
| 1 | 15 | 3.7122 | 1.9243 | 0.4969 | 1.4088 | 3.0348 |
| 2 | 13 | 1.8934 | 2.4531 | 0.6804 | 1.7591 | 4.0494 |

This report gives a confidence interval for the standard deviation in each group. Note that the accuracy of these intervals is very dependent on the assumption that the data were normally distributed.

## Confidence Interval for Standard Deviation Ratio

| Statistics | SD1 | SD2 | SD Ratio | 95% Lower Conf. Limit of Var. Ratio | 95% Upper Conf. Limit of Var. Ratio |
|---|---|---|---|---|---|
| SD Ratio | 1.9243 | 2.4531 | 0.784 | 0.192 | 1.877 |

This report gives a confidence interval for the ratio of the two standard deviations. Note that the accuracy of this interval is very dependent on the assumption that the data were normally distributed.

## Equal Variance Test

| Statistic | DF1 | DF2 | F-Value | Prob Level | Reject Hypothesis of Equal Variance at the 5.0% Significance Level? |
|---|---|---|---|---|---|
| Variance Ratio | 14 | 12 | 1.625 | 0.3834 | No |

This report provides a test of whether the two variances are equal. Unfortunately, when you fail to reject, you do not establish the validity of the equal variance assumption.

Chapter 210

# One-Way Analysis of Variance

## Introduction

This procedure performs a one-way (single-factor) analysis of variance and the Kruskal-Wallis one-way analysis of variance on ranks of data contained in either two or more variables or in one variable indexed by a second (grouping) variable. The one-way analysis of variance compares the means of two or more groups to determine if at least one group mean is different from the others. The F-ratio is used to determine statistical significance. The tests are nondirectional in that the null hypothesis specifies that all means are equal and the alternative hypothesis simply states that at least one mean is different.

## Kinds of Research Questions

One of the most common tasks in research is to compare two or more populations (groups). We might want to compare the income level of two regions, the nitrogen content of three lakes, or the effectiveness of four drugs. The first question that arises concerns which aspects (parameters) of the populations we should compare. We might consider comparing the means, medians, standard deviations, distributional shapes (histograms), or maximum values. We base the comparison parameter on our particular problem.

One of the simplest comparisons we can make is between the means of two or more populations. If we can show that the mean of one population is different from that of the other populations, we can conclude that the populations are different. Other aspects of the populations can (and should) also be considered, but the mean is usually the starting point.

If we are willing to make assumptions about other characteristics of the populations (such as that they are normally distributed and that their variances are equal), we can use the F-ratio to compare the means of random samples drawn from these populations. If these assumptions are violated, the nonparametric Kruskal-Wallis test may be used.

# Assumptions

The statistical tests described in this chapter make certain assumptions. One reason for the popularity of the F-test is its robustness in the face of assumption violation. However, if an assumption is not even approximately met, the significance levels and the power of the F-test are invalidated. Unfortunately, in practice it often happens that not one but several assumptions are not met. This makes matters even worse! Hence, steps should be taken to check the assumptions before important decisions are made. The reports include sections that investigate these assumptions.

## One-Way Analysis of Variance Assumptions

The assumptions of the **one-way analysis of variance** are:

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution. Each group is normally distributed about the group mean.
3. The variances of the populations are equal.
4. The groups are independent. There is no relationship among the individuals in one group as compared to another.
5. Each group is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

## Kruskal-Wallis Test Assumptions

The assumptions of the **Kruskal-Wallis test** are:

1. The variable of interest is continuous (not discrete). The measurement scale is at least ordinal.
2. The probability distributions of the populations are identical, except for location. Hence, we still require that the population variances are equal.
3. The groups are independent.
4. All groups are simple random samples from their respective populations. Each individual in the population has an equal probability of being selected in the sample.

# Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can assume that you have met the continuous variable assumption. Perhaps the greatest restriction is that your data come from a random sample of the population. If you do not have a random sample, your significance levels will be incorrect.

# Multiple Comparison Procedures

Given that the analysis of variance (ANOVA) test finds a significant difference among treatment means, the next task is to determine which treatments are different. Multiple comparison procedures (MCPs) are methods that pinpoint which treatments are different.

The discussion to follow considers the following experiment. Suppose an experiment studies how two gasoline additives influence the miles per gallon obtained. Three types of gasoline were studied. The first sample received additive W, the second received additive V, and the third did not receive an additive (the control group).

If the F-test from an ANOVA for this experiment is significant, we do not know which of the three possible pairs of groups are different. MCPs can help solve this dilemma.

## Multiple Comparison Considerations

Whenever MCPs are to be used, the researcher needs to contemplate the following issues.

### Exploration Versus Decision-Making

When conducting exploration (or data snooping), you make several comparisons to discover the underlying factors that influence the response. In this case, you do not have a set of planned comparisons to make. In contrast, in a decision-making mode, you would try to determine which treatment is preferred. In the above example, because you do not know which factors influence gasoline additive performance, you should use the exploration mode to identify those. A decision-making emphasis would choose the gasoline that provides the highest miles per gallon.

### Choosing a Comparison Procedure

You should consider two items here. First, will you know before or after experimentation which comparisons are of interest? Second, are you interested in some or all possible comparisons? Your choice of an MCP will depend on how you answer these two questions.

### Error Rates

You will need to consider two types of error rates: comparisonwise and experimentwise.

1. Comparisonwise error rate. In this case, you consider each comparison of the means as if it were the only test you were conducting. This is commonly denoted as $\alpha$. The conceptual unit is the comparison. Other tests that might be conducted are ignored during the calculation of the error rate. If we perform several tests, the probability of a type I error on each test is $\alpha$.
2. Experimentwise, or familywise, error rate. In this situation, the error rate relates to a group of independent tests. This is the probability of making one or more type I errors in a group of independent comparisons. We will denote this error rate as $\alpha_f$.

The relationship between these two error rates is:

$$\alpha_f = 1-(1-\alpha)^c$$

where *c* is the total number of comparisons in the family. The following table shows these error rates for a few values of *c* and $\alpha$. The body of the table consists of the calculated values of $\alpha_f$.

**Calculated Experimentwise Error Rates**

| | *c* | | | | |
|---|---|---|---|---|---|
| α | 2 | 3 | 5 | 10 | 20 |
| **0.20** | .360 | .488 | .672 | .893 | .988 |
| **0.10** | .190 | .271 | .410 | .651 | .878 |
| **0.05** | .098 | .143 | .226 | .401 | .642 |
| **0.02** | .040 | .059 | .096 | .183 | .332 |
| **0.01** | .020 | .030 | .049 | .096 | .182 |

As you can see, the possibility of at least one erroneous result goes up markedly as the number of tests increases. For example, in order to obtain an $\alpha_f$ of 0.05 with a *c* of 5, you would need to set $\alpha$ to 0.01.

## Multiple Comparison Procedure Definitions

All of the multiple comparison procedures (MCPs) considered here assume that there is independence between treatments or samples, equal variance for each treatment, and normality (except the Kruskal-Wallis Z, which does not need normality). In addition, unless stated otherwise, the significance tests are assumed to be two-tailed.

Let $\bar{y}_i$ and $n_i$ represent the mean and sample size of the i$^{th}$ treatment group. Let $s^2$ represent the the mean square error for these means based on *v* degrees of freedom. Let *k* be the number of treatments being compared for a factor or interaction.

### Alpha

This is the $\alpha_f$, or $\alpha$, specified for the multiple comparison test. It may be comparisonwise or experimentwise, depending on the test. This alpha can range from 0.01 to 0.10.

### Bonferroni (All Pairs)

The Bonferroni MCP chooses the comparisonwise error rate in such a way as to control the desired experimentwise $\alpha_f$. With k means and with an interest in all-possible pairs, the comparisonwise error rate is defined as $\alpha = \alpha_f / (k(k-1))$. The significance test for any pair would be as follows, where $t_{\alpha,v}$ is a Student's t with *v* degrees of freedom

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{s^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \geq t_{\alpha,v}$$

Generally, this MCP is run after the fact to find out which pairs are different.

## Bonferroni (Versus Control)

If one of the treatments is a control group and you want to compare all of the other means to the mean of this control group, there are k - 1 comparisons. Again, you should choose the comparisonwise error rate in such a way as to achieve the overall or experimentwise α. The comparisonwise error rate is $\alpha = \alpha_f /(2(k-1))$. The significance test for any two means would be as follows, where $t_{\alpha,v}$ is a Student's t

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{s^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \geq t_{\alpha,v}$$

## Comparison

This is a planned (a priori) significance test for a specific comparison that would also have a comparisonwise error rate. If you wanted to make several planned comparisons contained within one of the possible comparison options (the standard set of comparisons, the set of orthogonal polynomials, the set of comparisons with each treatment with the first treatment, the set with each treatment with the last treatment, or a set of no more than three customized contrasts), you could adjust the comparisonwise error rate to achieve a specific overall error rate. This test, distributed as a Student's $t_{\alpha/2,v}$, would be as follows, where $a_j$ are the comparison coefficients:

$$\frac{\left|\sum_j a_j \bar{y}_j\right|}{s\sqrt{\sum_j \frac{a_j^2}{n_j}}} \geq t_{\alpha/2,v}$$

## Duncan's

This MCP looks at all pairwise comparisons among k means, but the error rate is neither on an experimentwise nor on a comparisonwise basis. The error rate is based on the number of steps apart, *r*, the two means are when they are ordered**.** The probability of falsely rejecting the equality of two population means when the sample means are r steps apart is $1-(1-\alpha)^{r-1}$. The significance test is based on the Studentized range, $q_{\alpha,r,v}$:

$$\frac{|\bar{y}_i - \bar{y}_j|}{s/\sqrt{\frac{2}{\frac{1}{n_i} + \frac{1}{n_j}}}} \geq q_{\alpha,r,v}$$

## Dunnett's One and Two-Tailed Tests Versus a Control

If one of the treatments is a control group and you want to compare all of the other means to the mean of this control group, there are *k* - 1 comparisons. Dunnett's multiple comparison procedure (see Hsu(1996)) gives an experimentwise error rate of $\alpha$. The significance test for any two means would be as follows, where $q_{\alpha,v}$ is calculated using Dunnett's formulas.

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{s^2\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \geq q_{\alpha,v}$$

Often, it is of interest to find only those treatments that are better (or worse) than the control, so both one and two sided versions of this test are provided.

A set of two-sided, simultaneous confidence intervals are also provided for the difference between each treatment and the control.

## Fisher's LSD

Fisher's least significant difference (FSD) is a special version of the least significant difference (LSD). The difference between LSD and FSD is that FSD is only used when the F-test for the term is significant. LSD and FSD are used for pairwise comparisons.

The error rate for each comparison is comparisonwise. This test has no control of the experimentwise error rate. The significance test is as follows, where $\gamma = \alpha / 2$ for the LSD and $\gamma = \alpha / c$ for the FSD.

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{s^2\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \geq t_{\gamma,v}$$

## Hsu's Tests Versus the Best

Hsu (1996) chapter 4 provides a procedure for testing each group versus the best. This procedure is useful when you want to determine which of treatments is the best. Note that because of sampling variability, the sample best may not necessarily be the true best. Hsu's constrained multiple comparison with the best procedure allows the candidates for the best to be compared.

The method uses Dunnett's one-sided critical values, $q_i$, to provide simultaneous confidence intervals for

$$\mu_i - \max_{j \neq i}(\mu_j), i = 1, \cdots, k$$

which are constrained to include 0. The constraints were suggested by John W. Tukey because a confidence interval for the above quantity whose lower limit is 0 indicates that the $i^{th}$ treatment is the best. Likewise, a confidence interval for the above quantity whose upper limit is 0 indicates that the $i^{th}$ treatment is not the best.

Hsu's confidence intervals are given by

$$\left[D_i^-, D_i^+\right], i = 1, \cdots, k$$

where

$$D_i^+ = \max\left[0, \min_{j \neq i}\left(\left(\hat{\mu}_i - \hat{\mu}_j\right) + sq_i\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}\right)\right]$$

$$G = \left\{i : D_i^+ > 0\right\}$$

$$D_i^- = \begin{cases} 0 & if\ G = \{i\} \\ \min\limits_{j \in G, j \neq i}\left(\left(\hat{\mu}_i - \hat{\mu}_j\right) - sq_j\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}\right) & otherwise \end{cases}$$

## Kruskal-Wallis Z (Dunn's Test)

This test is attributed to Dunn (1964) and is referenced in Daniel (1990), pages 240 - 242. This MCP is a distribution-free multiple comparison, meaning that the assumption of normality is not necessary. It is to be used for testing pairs of medians following the Kruskal-Wallis test. The test needs sample sizes of at least five (but preferably larger) for each treatment. The error rate is adjusted on a comparisonwise basis to give the experimentwise error rate, $\alpha_f$. Instead of using means, this MCP uses average ranks, as the following formula indicates, with $\alpha = \alpha_f / \left(k\left(k-1\right)\right)$:

$$\frac{|\overline{R}_i - \overline{R}_j|}{\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \geq z_\alpha$$

Adjusted for ties the inequality becomes

$$\frac{|\overline{R}_i - \overline{R}_j|}{\sqrt{\frac{\left[N\left(N^2-1\right) - \left(\sum t^3 - \sum t\right)\right]\left[\frac{1}{n_i} + \frac{1}{n_j}\right]}{12(N-1)}}} \geq z_\alpha$$

In these inequalities, *N* is the total sample size and *t* is the number of values in the combined sample that are tied at a given rank.

### Newman-Keuls

The Newman-Keuls MCP relies on the number of ordered steps r, where r ranges from 2 to k, between two sample means. The error rate is neither experimentwise nor comparisonwise. Instead it is defined for sample means which are the same number of ordered steps apart. This test relies on the Studentized range distribution.

$$\frac{|\bar{y}_i - \bar{y}_j|}{s/\sqrt{\frac{2}{\frac{1}{n_i}+\frac{1}{n_j}}}} \geq q_{\alpha,r,v}$$

### Scheffe's

This MCP can be used to examine all possible comparisons among k means or just to look at all pairs as done here. It controls the overall or experimentwise error rate and is less sensitive than the Tukey-Kramer MCP. The significance test for pairs is as follows:

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{s^2\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \geq \sqrt{(k-1)F_{\alpha,k-1,v}}$$

### Tukey-Kramer

This test can be used to examine all pairs of treatment means. The error rate is experimentwise, and this test uses the Studentized range distribution. This test is conservative, which means that the two averages must be very different. The significance test follows:

$$\frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{\frac{s^2}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \geq q_{\alpha,k,v}$$

## Recommendations

These recommendations assume that normality and equal variance are valid. If normality is not valid for each treatment, then use the Kruskal-Wallis Z MCP.

1. Planned all-possible pairs. If you are interested in paired comparisons only and you know this in advance, use either the Bonferroni for pairs or the Tukey-Kramer MCP.
2. Unplanned all-possible pairs. Use Scheffe's MCP.
3. Each versus a control. Use Dunnett's test.
4. Selected but planned. Use Comparison and adjust the alpha level accordingly.
5. Comparison with the best. Use Hsu's procedure.

## Data Structure

The data may be entered in two formats, as shown in the examples below. The examples give the yield of corn for three types of fertilizer. The first format, shown in the first table below, puts the responses for each group in separate variables; that is, each variable contains all responses for a single group.

The second format, shown in the second table below, arranges the data so that all responses are entered in a single variable. A second variable, the Grouping Variable, contains an index that gives the group (A, B, or C) to which that row of data belongs.

In most cases, the second format is more flexible. Unless there is some special reason to use the first format, we recommend that you use the second.

**Three Response Variables**

| Yield A | Yield B | Yield C |
|---|---|---|
| 452 | 546 | 785 |
| 874 | 547 | 458 |
| 554 | 774 | 886 |
| 447 | 465 | 536 |
| 356 | 459 | |
| 754 | 665 | 669 |
| 558 | 467 | 857 |
| 574 | 365 | 821 |
| 664 | 589 | 772 |
| 682 | 534 | 732 |
| | 456 | 689 |
| 547 | 651 | 654 |
| | 654 | |
| 435 | 665 | 297 |
| | 546 | 830 |
| 245 | 537 | 827 |

**One Grouping and One Response Variable**

| Row | Fertilizer | Yield |
|---|---|---|
| 1 | B | 546 |
| 2 | B | 547 |
| 3 | B | 774 |
| 4 | B | 465 |
| 5 | B | 459 |
| 6 | B | 665 |
| 7 | B | 467 |
| 8 | B | 365 |
| 9 | B | 589 |
| 10 | B | 534 |
| 11 | B | 456 |
| 12 | B | 651 |
| 13 | B | 654 |
| 14 | B | 665 |
| 15 | B | 546 |
| 16 | B | 537 |
| 17 | A | 452 |
| 18 | A | 874 |
| 19 | A | 554 |
| 20 | A | 447 |
| 21 | A | 356 |

| Row | Fertilizer | Yield |
|---|---|---|
| 22 | A | 754 |
| 23 | A | 558 |
| 24 | A | 574 |
| 25 | A | 664 |
| 26 | A | 682 |
| 27 | A | 547 |
| 28 | A | 435 |
| 29 | A | 245 |
| 30 | C | 785 |
| 31 | C | 458 |
| 32 | C | 886 |
| 33 | C | 536 |
| 34 | C | 669 |
| 35 | C | 857 |
| 36 | C | 821 |
| 37 | C | 772 |
| 38 | C | 732 |
| 39 | C | 689 |
| 40 | C | 654 |
| 41 | C | 297 |
| | | 356 |

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

This panel specifies the variables used in the analysis.

### Response Variables

#### Response Variable(s)

This option lets you specify the variable(s) to be analyzed. Note that if you specify only one variable here, you must also specify a grouping variable. If you want to compare several variables (columns), you specify them here. If more than one variable is specified, only the variable numbers are displayed.

### Factor Variable

#### Factor Variable

The optional grouping (breakdown) variable indicates how the values of the response variable(s) should be grouped. Examples of grouping variables are males and females, age groups, “yes” or “no” responses, and so on. Note that the values in the variable may be either numeric or text. The treatment of text variables is specified for each variable by the Data Type option on the database.

A separate analysis is performed for each Response Variable when the Factor Variable is specified.

#### Type

This option specifies whether the factor is fixed or random. This is a formality in the one-way ANOVA since the F-test is identical no matter which option is selected. The selection influences the calculated power of the F-test as well as the expected mean squares.

### Planned Comparisons

#### Comparison

Specifies the planned comparisons that should be generated. Several predefined sets are available or you can specify up to three of your own in the Custom (1-3) options that follow. Each option will be explained next. Note that the contrasts are defined by a set of coefficients (see “Contrast” below).

- **None**

  This option indicates that no planned comparisons should be generated.

- **Standard Set**

  This option generates a standard (commonly used) set of contrasts. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -3,1,1,1 | Compare the first-level mean with the average of the rest. |
  | 0,-2,1,1 | Compare the second-level mean with the average of the rest. |
  | 0,0,-1,1 | Compare the third-level mean with the fourth-level mean. |

- **Orthogonal Polynomials**

  This option generates a set of orthogonal contrasts that allows you to test various trend components from linear up to sixth order. These contrasts are appropriate even if the levels are unequally spaced or the group sample sizes are unequal. Of course, these contrasts are only appropriate for data that is at least ordinal. Usually, you would augment the analysis of this type of data with a multiple regression analysis.

  The following example displays the type of contrast generated by this option. Suppose there are four equally spaced levels in the factor and each group has two observations. The contrasts generated by this option are (scaled to whole numbers):

  | | |
  |---|---|
  | -3,-1,1,3 | Linear component. |
  | 1,-1,-1,1 | Quadratic component. |
  | -1,3,-3,1 | Cubic component. |

- **Linear Trend**

  This option generates a set of orthogonal contrasts and retains only the linear component. This contrast is appropriate even if the levels are unequally spaced and the group sample sizes are unequal. See “Orthogonal Polynomials” above for more detail.

- **Linear-Quadratic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first two (the linear and quadratic) are reported.

- **Linear-Cubic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first three are reported.

- **Linear-Quartic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first four are reported.

- **Each with First**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the first level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,1,0,0 | Compare the first- and second-level means. |
  | -1,0,1,0 | Compare the first- and third-level means. |
  | -1,0,0,1 | Compare the first- and fourth-level means. |

- **Each with Last**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the last level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,0,0,1 | Compare the first- and fourth-level means. |
  | 0,-1,0,1 | Compare the second- and fourth-level means. |
  | 0,0,-1,1 | Compare the third- and fourth-level means. |

- **Custom**

  This option indicates that the contrasts listed in the next three boxes should be used.

## Planned Comparisons – Custom Comparisons

The following options are only used if Comparisons is set to 'Custom'.

### Custom (1-3)

This option lets you write a user-specified comparison by specifying the weights of that comparison. Note that there are no numerical restrictions on these coefficients. They do not even have to sum to zero. However, this is recommended. If the coefficients do sum to zero, the comparison is called a *contrast*. The significance tests anticipate that only one or two of these comparisons are to be run. If you run several, you should make some type of Bonferroni adjustment to your alpha value.

When you put in your own contrasts, you must be careful that you specify the appropriate number of weights. For example, if the factor has four levels, four weights must be specified, separated by commas. Extra weights are ignored. If too few weights are specified, the missing weights are assumed to be zero.

These comparison coefficients designate weighted averages of the level-means that are to be statistically tested. The null hypothesis is that the weighted average is zero. The alternative hypothesis is that the weighted average is nonzero. The weights (comparison coefficients) are specified here.

As an example, suppose you want to compare the average of the first two levels with the average of the last two levels in a six-level factor. You would enter “-1,-1,0,0,1,1.”

As a second example, suppose you want to compare the average of the first two levels with the average of the last three levels in a six-level factor. The contrast would be

-3,-3,0,2,2,2.

Note that in each case, we have used weights that sum to zero. This is why we could not use ones in the second example.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Reports

### Assumptions Report ... Means Report

Specify whether to display the indicated reports.

## Select Plots

### Means Plot and Box Plot

Specify whether to display the indicated plots.

## Report Options

### Test Alpha

The value of alpha for the statistical tests and power analysis. Usually, this number will range from 0.1 to 0.001. A common choice for alpha is 0.05, but this value is a legacy from the age before computers when only printed tables were available. You should determine a value appropriate for your particular study.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, whereas the double precision will display thirteen-place accuracy.

### Variable Names

Indicate whether to display the variable names or the variable labels.

### Value Labels

Indicate whether to display the data values or their labels.

## Multiple Comparison Tests

### Bonferroni (All-Pairs) ... Tukey-Kramer Confidence Intervals

These options specify which MC tests and confidence intervals to display.

## Multiple Comparison Tests – Options

### MC Alpha

Specifies the alpha value used by the multiple-comparison tests.

### MC Decimals

Specify how many decimals to display in the multiple comparison sections.

# Means Plot Tab

These options specify the plots of group means.

## Vertical and Horizontal Axis

### Label

This is the text of the axis labels. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

### Connect Lines

Click this box to connect the points for a particular factor. This makes it easier to spot patterns in the means.

### Titles

#### Plot Title

This option contains the text of the plot title. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

## Box Plot Tab

The options on this panel control the appearance of the box plot.

### Vertical and Horizontal Axis

#### Label

This is the text of the axis labels. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) axis. If left blank, these values are calculated from the data.

#### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on the vertical axis.

#### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

### Plot Settings

#### Plot Style File

Designate a box plot style file. This file sets all box plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Box Plot procedure.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by the appropriate variable names. Press the button on the right of the field to specify the font of the text.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running a One-Way ANOVA

This section presents an example of how to run a one-way analysis of variance. We will use the corn yield data contained in the SAMPLE database. These data are contained in the variables labeled YldA, YldB, and YldC.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the One-Way Analysis of Variance window.

**1 Open the SAMPLE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Sample.s0**.
- Click **Open**.

**2 Open the One-Way Analysis of Variance window.**

- On the menus, select **Analysis**, then **Analysis of Variance (ANOVA)**, then **One-Way Analysis of Variance**. The One-Way Analysis of Variance procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the One-Way Analysis of Variance window, select the **Variables tab**.
- Double-click in the **Response Variables** box. This will bring up the variable selection window.
- Select **YldA, YldB,** and **YldC** from the list of variables and then click **Ok**.
- Select **Custom** in the **Comparisons** list box.
- Enter **-2,1,1** in the **Custom 1** box.

**4 Specify the reports.**

- On the One-Way Analysis of Variance window, select the **Reports tab**.
- Check the **Duncan's Test** option of the Multiple Comparison Tests.
- Check the **Kruskal-Wallis Z Test (Dunn's Test)** option of the Multiple Comparison Tests.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports and charts will be displayed in the Output window.

## Tests of Assumptions Section

**Tests of Assumptions Section**

| Assumption | Test Value | Prob Level | Decision (0.05) |
|---|---|---|---|
| Skewness Normality of Residuals | -0.1787 | 0.858144 | Accept |
| Kurtosis Normality of Residuals | 0.4200 | 0.674472 | Accept |
| Omnibus Normality of Residuals | 0.2084 | 0.901062 | Accept |
| Modified-Levene Equal-Variance Test | 1.0866 | 0.347107 | Accept |

This section presents the results of tests validating the normality and equal variance assumptions. Note that the ANOVA assumes combined residuals (deviations for group means) are normal. Hence, the normality tests are performed on the combined set of residuals from all groups. Other assumptions concerning independence and random sampling are not tested here. You must justify those assumptions by considering your experiment procedure.

When using this report, all you need to do is scan down the column labeled *Decision(5%)*. If none of the tests are *rejected*, you can feel confident that the assumptions are met. (Of course, the power of these tests is also influenced by your sample size. If you have a small sample size, say less than 25 per group, the power of the normality tests will be questionable and you will have to rely on other means to justify your assumptions.)

Two aspects of normality are tested for, skewness and kurtosis. If the normality of residuals fails because of skewness, it might be possible to use the square root or logarithmic transformation to *normalize* your data.

Conover (1981) did extensive simulation involving different distributions, sample sizes, means, and variances; and they found that the modified-Levene test is one of the most robust and powerful tests for equality of variance. Thus, if a preliminary test is to be performed, use the modified-Levene test.

In the case of nonnormality, the Kruskal-Wallis nonparametric test is suggested. The basic assumptions of independent samples, continuous random variables, and a measurement scale of at least ordinal scale hold for this test. The Kruskal-Wallis test has the additional assumption that the distributions for the groups are identical (although not necessary normal) in form and shape (i.e., same variance) but differ only in location (i.e., in medians).

Finally, you should back up the results of these numerical tests by considering the box plots of the groups. As explained below, they let you visually determine if the assumptions of normality and equal variance are justified.

We next present the individual definitions of the items in this report.

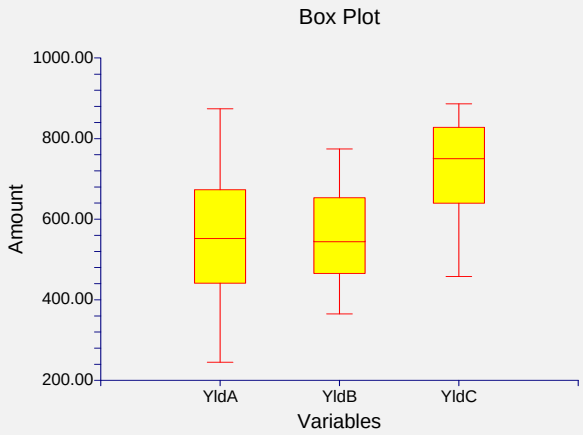
### Normality (Skewness, Kurtosis, and Omnibus)

These three tests allow you to test the skewness, kurtosis, and overall normality of the data. If any of them reject the hypothesis of normality, the data should not be considered normal. These tests are discussed in more detail in the "Descriptive Statistics" chapter.

### Equal-Variance Test (Modified Levene)

The modified Levene test has been found to be one of the best tests for equality of variances. The Levene (1960) procedure is outlined in the "Two-Sample Tests" chapter and will not be repeated here.

## Box Plots




Box plots are useful for assessing symmetry, presence of outliers, general equality of location, and equality of variation.

## Expected Mean Squares Section

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A( ... ) | 2 | Yes | S(A) | S+sA |
| S(A) | 40 | No | | S(A) |

Note: Expected Mean Squares are for the balanced cell-frequency case.

The Expected Mean Square expressions are provided to show the appropriate error term for each factor. The correct error term for a factor is that term that is identical except for the factor being tested.

### Source Term

The source of variation or term in the model.

### DF

The degrees of freedom. The number of observations "used" by this term.

### Term Fixed?

Indicates whether the term is “fixed” or “random.”

### Denominator Term

Indicates the term used as the denominator in the F-ratio.

### Expected Mean Square

This is the expected value of the mean square for the term in the ANOVA model assuming balanced data (equal group counts). *“S (A)”* represents the expected value of the mean square error (sigma). The uppercase letters represent either the adjusted sum of squared treatment means if the factor is fixed, or the variance component if the factor is random. The lowercase letter represents the number of levels for that factor, and “*s*” represents the number of replications of the whole experimental layout.

## Analysis of Variance Table Section

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A ( ... ) | 2 | 268532.4 | 134266.2 | 7.47 | .001746* | .925284 |
| S (Error) | 40 | 718574.3 | 17964.36 | | | |
| Total (Adjusted) | 42 | 987106.6 | | | | |
| Total | 43 | | | | | |

* Term significant at alpha = 0.05

### Source Term

The source of variation. The term in the model.

### DF

The degrees of freedom. The number of observations “used” by the corresponding model term.

### Sum of Squares

This is the sum of squares for this term. It is usually included in the ANOVA table for completeness, not for direct interpretation.

### Mean Square

An estimate of the variation accounted for by this term. The sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and the mean square of its corresponding error term. This is also called the F-test value.

### Prob Level

The significance level of the above F-ratio. The probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha level of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant. Note that if the value is significant at the specified value of alpha, a star is placed to the right of the F-Ratio.

### Power (Alpha=0.05)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error (β). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual differences among the population means.

The power value calculated here assumes that the population standard deviation is equal to the observed standard deviation and that the differences among the population means are exactly equal to the difference among the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

Generally, you would consider the power of the test when you accept the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If you accept the null hypothesis with high power, there is not much left to do. At least you know that the means are NOT different. However, if you accept the null hypothesis with low power, you can take one or more of the following actions:

1. Increase your alpha level. Perhaps you should be testing at alpha = .05 instead of alpha =.01. Increasing the alpha level will increase the power.
2. Increasing your sample size will increase the power of your test if you have low power. If you have high power, an increase in sample size will have little effect.
3. Decrease the magnitude of the variance. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

## Kruskal-Wallis One-Way ANOVA on Ranks

**Kruskal-Wallis One-Way ANOVA on Ranks**
**Hypotheses**
Ho: All medians are equal.
Ha: At least two medians are different.

**Test Results**

| Method | DF | Chi-Square (H) | Prob Level | Decision (0.05) |
|---|---|---|---|---|
| Not Corrected for Ties | 2 | 11.26741 | .003575 | Reject Ho |
| Corrected for Ties | 2 | 11.27082 | .003569 | Reject Ho |

| | |
|---|---|
| Number Sets of Ties | 4 |
| Multiplicity Factor | 24 |

**Group Detail**

| Group | Count | Sum of Ranks | Mean Rank | Z-Value | Median |
|---|---|---|---|---|---|
| YldA | 13 | 229.50 | 17.65 | -1.4941 | 554 |
| YldB | 16 | 279.00 | 17.44 | -1.8342 | 546 |
| YldC | 14 | 437.50 | 31.25 | 3.3564 | 752 |

This test is a nonparametric substitute for the one-way ANOVA when the assumption of normality is not valid. When in doubt about normality, play it safe and use this test. The assumptions for this test were given in the “Assumptions” section at the beginning of this chapter. Two key assumptions that we remind you of is that the distributions are at least ordinal in nature and that they are identical, except for location. This means that ties (repeated values) are

not acceptable. When ties are present in your data, you should use the corrected version of this test. We next present the individual definitions of items on this report.

### Hypotheses

The null hypothesis is that the medians are equal versus the alternative that at least one median is different from the rest.

### Method

The results of two tests are presented. The first line gives the Kruskal-Wallis test with no correction for ties. The second line reports a modified Kruskal-Wallis test that has been modified to adjust for ties. If there are no ties, the results are identical.

### DF

The degrees of freedom of the large sample Chi-square approximation to the Kruskal-Wallis test distribution. Note that the degrees of freedom are equal to the number of groups minus one.

### Chi-Square (H)

The value of H, the uncorrected (for ties) Kruskal-Wallis test statistic. The formula for H is

$$H = \frac{12}{N\,(N\,+1)} \sum_{i=1}^{k} \frac{R_i^{\,2}}{n_i} - 3(N\,+1)$$

The Kruskal-Wallis test corrected for ties is calculated by dividing H by a correction factor. The formula for the corrected version of H is

$$H_C = \frac{H}{1 - \dfrac{\sum t(t^{\,2}-1)}{N\,(N^{\,2}-1)}}$$

In both of the above formulas, $\boldsymbol{N}$ is the total sample size, $\boldsymbol{n_i}$ is the sample size of the i$^{th}$ group, $\boldsymbol{k}$ is the number of groups, $\boldsymbol{R_i}$ is the sum of the ranks of the i$^{th}$ group, and $\boldsymbol{t}$ is the count of a particular tie.

### Prob Level

The significance level of H assuming a Chi-square distribution. The probability of an H larger than that obtained by this analysis. For example, to test at an alpha level of 0.05, this probability would have to be less than 0.05 to make H significant.

### Decision(0.05)

The decision about the null hypothesis based on this test.

### Number Sets of Ties

This is the number of sets of tied values. If there are no ties, this number is zero. A set of ties is two or more observations with the same value.

### Multiplicity Factor

This is the tie portion of the correction factor for H.

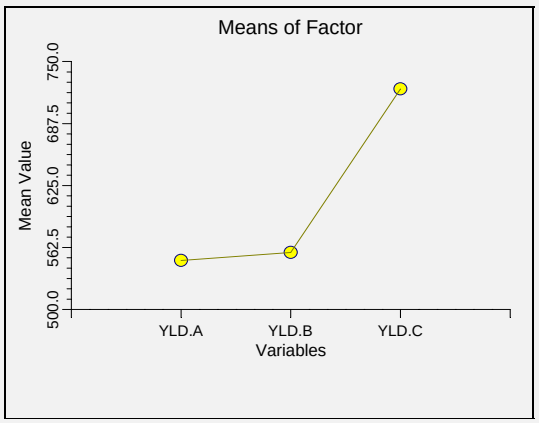
$$\sum_{j=1} (t_j^{\,3} - t_j)$$

# Means, Effects, and Plots Section

**Means and Effects Section**

| Term | Count | Mean | Standard Error | Effect |
|---|---|---|---|---|
| All | 43 | 608.7209 | | 609.7473 |
| A: | | | | |
| YldA | 13 | 549.3846 | 37.17356 | -60.36264 |
| YldB | 16 | 557.5 | 33.50779 | -52.24725 |
| YldC | 14 | 722.3571 | 35.82134 | 112.6099 |

**Plots Section**




### Term

The label for this line of the report.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Standard Error

The standard error of the mean. Note that the standard errors are simply the square root of the mean square of the error term for this term divided by the count. These standard errors are not the same as the simple standard errors calculated separately for each group. The standard errors reported here are those appropriate for use in testing multiple comparisons.

### Effect

The component that this term contributes to the mean. For example, the mean of the first group is equal to the sum of the overall effect (from the “All” line) plus the effect of the first term.

### Plot of Means

This plot displays the means for the data analyzed. Note how easily you can see patterns in the plot.

## Multiple-Comparison Sections

**Duncan's Multiple-Comparison Test**

Response: YldA,YldB,YldC
Term A:

Alpha=0.050 Error Term=S (A) DF=40 MSE=17964.36

| Group | Count | Mean | Different From Groups |
|---|---|---|---|
| YldA | 13 | 549.3846 | YldC |
| YldB | 16 | 557.5 | YldC |
| YldC | 14 | 722.3571 | YldA, YldB |

This section presents the results of the multiple-comparison procedures selected. These reports all use a uniform format that will be described by considering Duncan's Multiple-Comparison Test. The reports for the other procedures are similar. For more information on the interpretation of the various multiple-comparison procedures, turn to the section by that name.

We next present the individual definitions of items on this report.

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom of the error term.

### MSE

The value of the mean square error.

### Group

The label for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Different From Groups

A list of those groups that are significantly different from this group according to this multiple-comparison procedure. All groups not listed are not significantly different from this group.

## Planned-Comparison Section

This section presents the results of any planned comparisons that were selected.

**Planned Comparison: A1**

Response: YldA,YldB,YldC
Term A:

Alpha=0.050 Error Term=S (A) DF=40 MSE=17964.36

Comparison Value=181.0879 T-Value=2.0331 Prob>|T|=.048716 Decision(0.05)=Reject
Comparison Std Error=89.06983

| Group | Comparison Coefficient | Count | Mean |
|---|---|---|---|
| YldA | -2 | 13 | 549.3846 |
| YldB | 1 | 16 | 557.5 |
| YldC | 1 | 14 | 722.3571 |

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom of the error term.

### MSE

The value of the mean square error.

### Comparison Value

The value of the comparison. This is formed by the multiplying the Comparison Coefficient times the Mean for each group and summing.

### T-Value

The t-test used to test whether the above Comparison Value is significantly different from zero.

$$t_f = \frac{\sum_{i=1}^{k} c_i M_i}{\sqrt{MSE \sum_{i=1}^{k} \frac{c_i^2}{n_i}}}$$

where ***MSE*** is the mean square error, ***f*** is the degrees of freedom associated with ***MSE***, ***k*** is the number of groups, $c_i$ is the comparison coefficient for the i$^{th}$ group, $M_i$ is the mean of the i$^{th}$ group, and $n_i$ is the sample size of the i$^{th}$ group.

### Prob>|T|

The significance level of the above T-Value. The Comparison is statistically significant if this value is less than the specified alpha.

### Decision(0.05)

The decision based on the specified value of the multiple-comparison alpha.

### Comparison Standard Error

This is the standard error of the estimated comparison value. It is the denominator of the T-Value (above).

### Group

The label for this group.

### Comparison Coefficient

The coefficient (weight) used for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

## Kruskal-Wallis Multiple-Comparison Z-Value Section

**Kruskal-Wallis Multiple-Comparison Z-Value Test (Dunn's Test)**

| Variable | YldA | YldB | YldC |
|---|---|---|---|
| YldA | 0.0000 | 0.0462 | 2.8117 |
| YldB | 0.0462 | 0.0000 | 3.0063 |
| YldC | 2.8117 | 3.0063 | 0.0000 |

Regular Test: Medians significantly different if z-value > 1.9600
Bonferroni Test: Medians significantly different if z-value > 2.3940

### Z-Values

The values in the table are appropriate for testing whether the medians of any two groups are significantly different. The formula for $z_{ij}$ (comparing group i to group j) is

$$z_{ij} = \frac{\left|\frac{R_i}{n_i} - \frac{R_j}{n_j}\right|}{\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

In the presence of ties, the adjusted formula is

$$z_{ij} = \frac{\left|\frac{R_i}{n_i} - \frac{R_j}{n_j}\right|}{\sqrt{\frac{\left[N\left(N^2-1\right) - \left(\sum t^3 - \sum t\right)\right]\left[\frac{1}{n_i} + \frac{1}{n_j}\right]}{12(N-1)}}}$$

where $\boldsymbol{N}$ is the total sample size, $\boldsymbol{n_i}$ is the sample size of the i$^{th}$ group, $\boldsymbol{t}$ is the number of values in the combined sample that are tied at a given rank, and $\boldsymbol{R_i}$ is the sum of the ranks of the i$^{th}$ group.

The distribution of $z_{ij}$ is normal with mean equal to zero and variance equal to one. If you are only making one or two tests, you would compare the value in the table to the Regular Test value, $z_{\alpha/2}$. If the computed $z_{ij}$ is greater than this value, the two groups are significantly different.

However, if you are using all the tests from the table, you should use the Bonferroni Test value. This is a *z*-value that has been adjusted for multiple tests by dividing $\alpha / 2$ by *k*(*k*-1)/2, making it $z_{\alpha/(k(k-1))}$. Note the *k*(*k*-1)/2 is the number of possible pairs of *k* groups.

If you are making a specific number of tests, say m, that is less than all-possible pairs, you will have to manually make the correct adjustment by dividing $\alpha / 2$ by *m*. This might happen if you are comparing each treatment group with a control group, in which case you would have *k* - 1 tests.

# One-Way ANOVA Checklist

This checklist, prepared by a professional statistician, is a flowchart of the steps you should complete to conduct a valid one-way ANOVA (or its nonparametric counterpart). You should complete these tasks in order.

## Step 1 – Data Preparation

### Introduction

This step involves scanning your data for anomalies, data entry errors, typos, and so on.

### Sample Size

The sample size (number of nonmissing rows) has a lot of ramifications. The one-way ANOVA was developed under the assumption that the sample sizes in each group are equal. In practice, this seldom happens, but the closer you can get to equal sample sizes the better.

With regard to the combined sample size, the ANOVA may be performed on very small samples, such as 4 or 5 observations per group. However, in order to test assumptions and obtain reliable estimates of variation, you should attempt to obtain at least 30 individuals per group.

### Missing Values

The number and pattern of missing values are always issues to consider. Usually, we assume that missing values occur at random throughout your data. If this is not true, your results will be biased since a particular segment of the population is underrepresented. If you have a lot of missing values, some researchers recommend comparing other variables with respect to missing versus nonmissing. If you find large differences in other variables, you should begin to worry about whether the missing values are cause for a systematic bias in your results.

### Type of Data

The mathematical basis of the F-test assumes that the data are continuous. Because of the rounding that occurs when data are recorded, all data are technically discrete. The validity of the assumption of the continuity of the data then comes down to determining when we have too much rounding. For example, most statisticians would not worry about human-age data that were rounded to the nearest year. However, if these data were rounded to the nearest ten years or further to only three groups (young, adolescent, and adult), most statisticians question the validity of the probability statements. Some studies have shown that the F-test is reasonably accurate when the data have only five possible values (most would call this discrete data). If your data contain fewer than five unique values, any probability statements made are tenuous.

### Outliers

Generally, outliers cause distortion in F-tests. You must scan your data for outliers (the box plot is an excellent tool for doing this). If you have outliers, you have to decide if they are one-time occurrences or if they would occur in another sample. If they are one-time occurrences, you can remove them and proceed. If you know they represent a certain segment of the population, you have to decide between biasing your results (by removing them) or using a nonparametric test that can deal with them. Most would choose the nonparametric test.

## Step 2 – Setup and Run the Panel

### Introduction

Now comes the fun part: running the program. **NCSS** is designed to be simple to operate, but it can still seem complicated. Before you run a procedure such as this for the first time, take a few minutes to read through the chapter again and familiarize yourself with the issues involved.

### Enter Variables

The ***NCSS*** procedures were set with ready-to-run defaults. About all you have to do is select the appropriate variables (columns of data).

### Select All Plots

As a rule, you should select all diagnostic plots, even though they may take a few extra seconds to generate. They add a great deal to your analysis of the data.

### Specify Alpha

Most beginners at statistics forget this important step and let the alpha value default to the standard 0.05. You should make a conscious decision as to what value of alpha is appropriate for your study. The 0.05 default came about during the dark ages when people had to rely on printed probability tables and there were only two values available: 0.05 or 0.01. Now you can set the value to whatever is appropriate.

A special note on setting the Multiple Comparison alpha. We suggest that you set this at 0.10 so that the individual tests are made at a more reasonable significance level.

# Step 3 – Check Assumptions

## Introduction

Once the output is displayed, you will be tempted to go directly to the probability of the F-test, determine if you have a significant result, and proceed to something else. However, it is very important that you proceed through the output in an orderly fashion. The first task is to determine if the assumptions are met by your data.

Sometimes, when the data are nonnormal for all samples, a data transformation (like square roots or logs) might normalize the data. Frequently, when one sample is normal and the other is not, this transformation, or re-expression, approach works well.

## Random Sample

The validity of this assumption depends upon the method used to select the sample. If the method used assures that each individual in the population of interest has an equal probability of being selected for this sample, you have a random sample. Unfortunately, you cannot tell if a sample is random by looking at it or statistics from it.

## Sample Independence

The samples must be independent. For example, if you randomly divide a sample of individuals into two groups, you have met this requirement. However, if your population consists of cars and you assign the left tire to one group and the right tire to the other, you do not have independence. Here again, you cannot tell if the samples are independent by looking at them. You must consider the sampling methodology.

## Check Means Report

You should check the Means and Effects Section first to determine if the Counts and the Means are reasonable. If you have selected the wrong variable, these values will alert you.

## Normality

To validate this assumption, you should first look at the plots. Outliers will show up on the box plots and the probability plots. No data will be perfectly normal. After considering the plots, look at the Tests of Assumptions section to get numerical confirmation of what you see in the plots. Remember that the power of these normality tests is directly related to the sample size, so when the normality assumption is accepted, double check that your sample is large enough to give conclusive results.

## Equal Variance

The equal variance assumption is important in determining which statistical test to use. Check the box plots for boxes with about the same widths. Confirm your conclusion by looking at the Equal-Variance Test (Modified Levene) line.

## Step 4 – Choose the Appropriate Statistical Test

### Introduction

You are now ready to determine which statistical procedures will be valid.

### Normal Data with Equal Variances

Use the Analysis of Variance Section for hypothesis testing.

### Normal Data with Unequal Variances

Try variance stabilizing transformations like the log or square root. If this does not work, you might try testing two groups at a time using the unequal variance two-sample t-tests. If you decide to make several t-tests, you should make appropriate adjustments to your significance level to avoid the multiplicity problem discussed in the Multiple Comparison section. The Kruskal-Wallis tests assumes that the variances are equal, so it cannot be used.

### Nonnormal Data with Equal Variances

Use the Kruskal-Wallis Test for hypothesis testing.

### Nonnormal Data with Unequal Variances

If you cannot find a variance-stabilizing transformation, you might test each pair of groups using the Kolmogorov-Smirnov test. Of course, the Kolmogorov-Smirnov test tests both the mean and variance. Since you already know that the variances are different from the Levene test, it is questionable whether this test will add new information. If you decide to make several Kolmogorov-Smirnov tests, you should make appropriate adjustments to your significance level to avoid the multiplicity problem discussed in the Multiple Comparison section.

## Step 5 – Interpret Findings

### Hypothesis Testing

The interpretation of an analysis of variance table is rather easy. You simply look at the *Prob>F* value. If this value is less than your chosen significance level (say .05), you can declare that at least two of the means are significantly different. You then determine which means are different using planned comparisons or an appropriate paired-comparison procedure. With a list of significantly different means, you can view the plot of the means and discuss the meaning of your results.

## Step 6 – Record Your Results

Finally, as you finish a test, take a moment to jot down what decisions you made and what you have found. Explain what you did, why you did it, what conclusions you reached, which outliers you deleted, areas for further investigation, and so on.

Chapter 211

# Analysis of Variance for Balanced Data

## Introduction

This procedure performs an analysis of variance on up to ten factors. The experimental design must be of the factorial type (no nested or repeated-measures factors) with no missing cells. If the data are balanced (equal-cell frequency), this procedure yields exact F-tests. If the data are not balanced, approximate F-tests are generated using the method of unweighted means (UWM).

The F-ratio is used to determine statistical significance. The tests are nondirectional in that the null hypothesis specifies that all means for a specified main effect or interaction are equal and the alternative hypothesis simply states that at least one is different.

Studies have shown that the properties of UWM F-tests are very good if the amount of unbalance in the cell frequencies is small. Despite that relative accuracy, you might well ask, “If the results are not always exact, why provide the method?” The answer is that the general linear models (GLM) solution (discussed in the General Linear Models chapter) sometimes requires more computer time and memory than is available. When there are several factors each with many levels, the GLM solution may not be obtainable. In these cases, UWM provides a very useful approximation. When the design is balanced, both procedures yield the same results, but the UWM method is much faster.

The procedure also calculates Friedman’s two-way analysis of variance by ranks. This test is the nonparametric analog of the F-test in a randomized block design. (See Help File for details.)

## Kinds of Research Questions

A large amount of research consists of studying the influence of a set of independent variables on a response (dependent) variable. Many experiments are designed to look at the influence of a single independent variable (factor) while holding other factors constant. These experiments are called single-factor experiments and are analyzed with the one-way analysis of variance (ANOVA). A second type of design considers the impact of one factor across several values of other factors. This experimental design is called the factorial design.

The factorial design is popular among researchers because it not only lets you study the individual effects of several factors in a single experiment, but it also lets you study their

interaction. Interaction is present when the response variable fails to behave the same at values of one factor when a second factor is varied. Since factors seldom work independently, the study of their interaction becomes very important.

## The Linear Model

We begin with an infinite population of individuals with many measurable characteristics. These individuals are separated into two or more treatment populations based on one or more of these characteristics. A random sample of the individuals in each population is drawn. A treatment is applied to each individual in the sample and an outcome is measured. The data so obtained are analyzed using an analysis of variance table, which produces an F-test.

A mathematical model may be formulated that underlies each analysis of variance. This model expresses the response variable as the sum of parameters of the population. For example, a linear mathematical model for a two-factor experiment is

$$Y_{ijk} = m + a_i + b_j + (ab)_{ij} + e_{ijk}$$

where i=1,2,...,I; j=1,2,...,J; and k=1,2,...,K. This model expresses the value of the response variable, *Y*, as the sum of five components:

$m$ the mean.

$a_i$ the contribution of the $i^{th}$ level of a factor A.

$b_j$ the contribution of the $j^{th}$ level of a factor B.

$(ab)_{ij}$ the combined contribution of the $i^{th}$ level of a factor A and the $j^{th}$ level of a factor B.

$e_{ijk}$ the contribution of the $k^{th}$ individual. This is often called the “error.”

Note that this model is the sum of various constants. This type of model is called a linear model and becomes the mathematical basis for our discussion of the analysis of variance. Also note that this serves only as an example. Many linear models could be formulated for the two-factor experiment.

## Assumptions

The following assumptions are made when using the F-test:

1. The response variable is continuous.
2. The $e_{ijk}$ follow the normal probability distribution with mean equal to zero.
3. The variances of the $e_{ijk}$ are equal for all values of i, j, and k.
4. The individuals are independent.

## Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can assume that you have met the continuous variable assumption. Perhaps the greatest restriction is that your data comes

from a random sample of the population. If you do not have a random sample, the F-test will not work.

The UWM procedure also requires that there are no missing cells. Because the concept of missing cells often gets confused with unbalanced data, we will give an example that discriminates between these two properties.

Let's assume that an experiment is designed to study the impact of education and region on income. Three regions are selected for this study. They are Boston, Chicago, and Denver. Two education levels are selected: high school and college. Hence, the experiment is a two-by-three factorial design with six treatment combinations (called "cells"). Suppose the researcher intends to sample ten individuals from each of the six treatment groups. If the experiment proceeds as planned, it will be balanced with no missing cells.

As long as there are ten individuals in each of the six cells, the design is said to be "balanced." Suppose that for one reason or another, two of the ten college people are lost from the Denver-college group. The design is now "unbalanced." Hence, an unbalanced design is one which has a differing number of individuals in each treatment group.

Suppose that instead of just two people, all ten individuals in the Denver-college group (cell) are lost from the study. Now the design has a missing cell." That is, one complete treatment combination is missing.

Again, the UWM procedure is exact for a balanced design, approximate for an unbalanced design with no missing cells, and impossible for a design with missing cells. Unfortunately, designs that are confounded, such as Latin squares and fractional factorials, have missing cells, so they cannot be analyzed with this procedure.

## Multiple Comparison Procedures

The multiple comparison procedures are discussed in the One-Way Analysis of Variance chapter.

## Friedman's Rank Test

Friedman's test is a nonparametric analysis that may be performed on data from a randomized block experiment. In this test, the original data are replaced by their ranks. It is used when the assumptions of normality and equal variance are suspect. In a experiment with $b$ blocks and $k$ treatments, the Friedman test statistic, $Q$, is calculated as follows:

$$Q = (k-1)\frac{12\sum_{j=1}^{k} R_j^2 - 3b^2k(k+1)^2}{bk(k^2-1) - \sum(t^3 - t)}$$

The data within each of the $b$ blocks are assigned ranks. The ranks are summed for each of the $k$ groups. This rank sum is denoted as $R_j$. The factor involving $t$ in the denominator is a correction for ties. The symbol $t$ represents the number of times a single value is found within a block. When the *multiplicity* $\sum(t^3 - t)$ is included, the test is said to be corrected for ties. When this term is omitted, the test value is not corrected for ties.

This statistic is approximately distributed as a Chi-square with *k-1* degrees of freedom.

The *Q* statistic is closely related to Kendall's coefficient of concordance, *W*, using the formula:

$$W = Q / b(k - 1)$$

In order to run this procedure, the first factor must be the blocking (random) factor and the second must be the treatment (fixed) factor.

# Data Structure

The data must be entered in a format that puts the response in one variable and the values of each of the factors in other variables. An example of the data for a randomized-block design is shown next.

**RNDBLOCK dataset**

| Block | Treatment | Response |
|---|---|---|
| 1 | 1 | 123 |
| 2 | 1 | 230 |
| 3 | 1 | 279 |
| 1 | 2 | 245 |
| 2 | 2 | 283 |
| 3 | 2 | 245 |
| 1 | 3 | 182 |
| 2 | 3 | 252 |
| 3 | 3 | 280 |
| 1 | 4 | 203 |
| 2 | 4 | 204 |
| 3 | 4 | 227 |

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

These panels specify the variables used in the analysis.

### Response Variables

#### Response Variable(s)

Specifies the response (dependent) variable to be analyzed. If you specify more than one variable here, a separate analysis is run for each variable.

### Factor Specification

#### Factor Variable

At least one factor variable must be specified. This variable's values indicates how the values of the response variable should be categorized. Examples of factor variables are gender, age groups,

"yes" or "no" responses, and the like. Note that the values in the variable may be either numeric or text. The treatment of text variables is specified for each variable by the Data Type option on the database.

### Type

This option specifies whether the factor is fixed or random. The selection influences the expected mean square, which in turn determines the denominator of the F-test.

A fixed factor includes all possible levels, like male and female for gender, includes representative values across the possible range of values, like low, medium, and high temperatures, or includes a set of vlaues to which inferences will be limited, like New York, California, and Maryland.

A random factor is one in which the chosen levels represent a random sample from the population of values. For example, you might select four classes from the hundreds in your state or you might select ten batches from an industrial process. The key is that a random sample is chosen.

### Comparisons

Comparisons are only generated for fixed factors. These options let you specify any planned comparisons that you want to run on this factor. A planned comparison is formulated in terms of the means as follows:

$$C_i = \sum_{j=1}^{J} w_{ij} m_j$$

In this equation, there are J levels in the factor, the means for each level of the factor are denoted $m_j$, and $w_{ij}$ represents a set of J weight values for the $i^{th}$ comparison. The comparison value, $C_i$, is tested using a t-test. Note that if the $w_{ij}$ sum to zero across j, the comparison is called a "contrast" of the means.

Comparisons may be specified by simply listing the weights. For example, suppose a factor has three levels (unique values). Further suppose that the first level represents a control group, the second a treatment at one dose, and the third a treatment at a higher dose. Three comparisons come to mind: compare each of the treatment groups of the control group and compare the two treatment groups to each other. These three comparisons would be

| | |
|---|---|
| Control vs. Treatment 1 | -1,1,0 |
| Control vs. Treatment 2 | -1,0,1 |
| Treatment 1 vs. Treatment 2 | 0,-1,1 |

You might also be interested in comparing the control group with the average of both treatment groups. The weights for this comparison would be -2,1,1.

When a factor is quantitative, it might be of interest to divide the response pattern into linear, quadratic, cubic, and similar components. If the sample sizes are equal and the factor levels are equally spaced, these so-called components of trend may be studied by the use of simple contrasts. For example, suppose a quantitative factor has three levels: 5, 10, and 15. A contrast to test the linear and quadratic trend components would be

| | |
|---|---|
| Linear trend | -1,0,1 |
| Quadratic trend | 1,-2,1 |

If the sample sizes for the groups are unequal (the design is unbalanced), adjustments must be made for the differing sample sizes.

***NCSS*** will automatically generate some of the more common sets of contrasts or it will let you specify up to three custom contrasts yourself. The following common sets are designated by this option.

- **None**

  No comparisons are generated.

- **Standard Set**

  This option generates a standard set of contrasts in which the mean of the first level is compared to the average of the rest, the mean of the second group is compared to the average of those remaining, and so on.

  The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -3,1,1,1 | Compare the first-level mean with the average of the rest. |
  | 0,-2,1,1 | Compare the second-level mean with the average of the rest. |
  | 0,0,-1,1 | Compare the third-level mean with the fourth-level mean. |

- **Polynomial**

  This option generates a set of orthogonal contrasts that allow you to test various trend components from linear up to sixth order. These contrasts are appropriate even if the levels are unequally spaced or the group sample sizes are unequal. Of course, these contrasts are only appropriate for data that are at least ordinal. Usually, you would augment the analysis of this type of data with a multiple regression analysis.

  The following example displays the type of contrast generated by this option. Suppose there are four equally spaced levels in the factor and each group has two observations. The contrasts generated by this option are (scaled to whole numbers):

  | | |
  |---|---|
  | -3,-1,1,3 | Linear component. |
  | 1,-1,-1,1 | Quadratic component. |
  | -1,3,-3,1 | Cubic component. |

- **Linear Trend**

  This option generates a set of orthogonal contrasts and retains only the linear component. This contrast is appropriate even if the levels are unequally spaced and the group sample sizes are unequal. See Orthogonal Polynomials above for more detail.

- **Linear-Quadratic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first two (the linear and quadratic) are reported.

- **Linear-Cubic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first three are reported.

- **Linear-Quartic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first four are reported.

- **Each with First**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the first level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,1,0,0 | Compare the first- and second-level means. |
  | -1,0,1,0 | Compare the first- and third-level means. |
  | -1,0,0,1 | Compare the first- and fourth-level means. |

- **Each with Last**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the last level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,0,0,1 | Compare the first- and fourth-level means. |
  | 0,-1,0,1 | Compare the second- and fourth-level means. |
  | 0,0,-1,1 | Compare the third- and fourth-level means. |

- **Custom**

  This option indicates that the contrasts listed in the corresponding three boxes of the Comparison panel should be used.

## Custom Comparisons Tab

This panel is used when the Comparison option of one or more factors is set to Custom. The Custom option means that the contrast coefficients are to be entered by the user. The boxes on this panel contain the user-supplied contrast coefficients. The first row is for factor one, the second row for factor two, and so on.

### Custom Comparisons

The following options are only used if Comparisons is set to 'Custom' on the Variables tab.

#### Custom (1-3)

This option lets you write a user-specified comparison by specifying the weights of that comparison. Note that there are no numerical restrictions on these coefficients. They do not even have to sum to zero; however, this is recommended. If the coefficients do sum to zero, the comparison is called a “contrast.” The significance tests anticipate that only one or two of these comparisons are to be run. If you run several, you should make some type of Bonferroni adjustment to your alpha value.

When you put in your own contrasts, you must be careful that you specify the appropriate number of weights. For example, if the factor has four levels, four weights must be specified, separated by commas. Extra weights are ignored. If not enough weights are specified, they are assumed to be zero.

These comparison coefficients designate weighted averages of the level-means that are to be statistically tested. The null hypothesis is that the weighted average is zero. The alternative

hypothesis is that the weighted average is nonzero. The weights (comparison coefficients) are specified here.

As an example, suppose you want to compare the average of the first two levels with the average of the last two levels in a six-level factor. You would enter “-1,-1,0,0,1,1.”

As a second example, suppose you want to compare the average of the first two levels with the average of the last three levels in a six-level factor. The contrast would be

-3,-3,0,2,2,2.

Note that in each case, we have used weights that sum to zero. This is why we could not use ones in the second example.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Reports

### EMS Report ... Means Report

Specify whether to display the indicated reports and plots.

## Select Plots

### Means Plot(s)

Specify whether to display the indicated plots.

## Report Options

### Test Alpha

The value of alpha for the statistical tests and power analysis. Usually, this number will range from 0.10 to 0.001. A common choice for alpha is 0.05, but this value is a legacy from the age before computers when only printed tables were available. You should determine a value appropriate for your particular study.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, while the double precision will display thirteen-place accuracy.

### Variable Names

Indicate whether to display the variable names or the variable labels.

### Value Labels

Indicate whether to display the data values or their labels.

## Multiple Comparison Tests

### Bonferroni Test (All-Pairs) ... Tukey-Kramer Confidence Intervals

These options specify which MC tests and confidence intervals to display.

### Tests for Two-Factor Interactions

This option specifies whether multiple comparison tests are generated for two-factor interaction terms. When checked, the means of two-factor interactions will be tested by each active multiple comparison test. The multiple comparison test will treat the means as if they came from a single factor. For example, suppose factor A as two levels and factor B has three levels. The AB interaction would then have six levels. The active multiple comparison tests would be run on these six means.

Care must be used when interpreting multiple comparison tests on interaction means. Remember that the these means contain not only the effects of the interaction, but also the main effects of the two factors. Hence these means contain the combined effects of factor A, factor B, and the AB interaction. You cannot interpret the results as representing only the AB interaction.

## Multiple Comparison Tests – Options

### MC Alpha

Specifies the alpha value used by the multiple-comparison tests.

### MC Decimals

Specify how many decimals to display in the multiple comparison sections.

# Means Plot Tab

These options specify the plots of group means.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Y Scaling

Specify the method for calculating the minimum and maximum along the vertical axis. *Separately* means that each plot is scaled independently. *Uniform* means that all plots use the overall minimum and maximum of the data. This option is ignored if a minimum or maximum is specified.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Lines

Click this box to connect the points for a particular factor. This makes it easier to spot patterns in the means.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, {X} is replaced by the appropriate factor name.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

# Symbols Tab

These options specify the symbols used in the plots of group means.

## Plotting Symbols

### Group (1-15)

The symbols used to represent the levels of a factor on the means plots. Group 1 represents the first level, Group 2 represents the second level, and so on.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Analysis of a Randomized-Block Design

This section presents an example of how to run an analysis of the data contained in the RNDBLOCK database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Analysis of Variance for Balanced Data window.

**1 Open the RNDBLOCK dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **RNDBLOCK.s0**.
- Click **Open**.

**2 Open the Analysis of Variance for Balanced Data window.**

- On the menus, select **Analysis**, then **Analysis of Variance (ANOVA)**, then **Analysis of Variance for Balanced Data**. The Analysis of Variance for Balanced Data procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Analysis of Variance for Balanced Data window, select the **Variables tab**.
- Double-click in the **Response Variables** box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**.
- Double-click in the **Factor 1 Variable** box. This will bring up the variable selection window.
- Select **Block** from the list of variables and then click **Ok**.
- Select **Random** in the **Type** box for Factor 1.
- Double-click in the **Factor 2 Variable** box. This will bring up the variable selection window.
- Select **Treatment** from the list of variables and then click **Ok**.
- Select **Fixed** in the **Type** box for Factor 2.
- Select **Linear** in the **Comparisons** box for Factor 2.

**4 Specify the reports.**

- On the Analysis of Variance for Balanced Data window, select the **Reports tab**.
- Check the **Tukey-Kramer Test** option of the Multiple Comparison Tests.

### 5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

We will now document this output, one section at a time.

## Expected Mean Squares Section

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (Block) | 2 | No | S | S+bsA |
| B (Treatment) | 3 | Yes | AB | S+sAB+asB |
| AB | 6 | No | S | S+sAB |
| S | 0 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

The Expected Mean Squares expressions are provided to show the appropriate error term for each factor. The correct error term for a factor is that term that is identical except for the factor being tested.

### Source Term

The source of variation or term in the model.

### DF

The degrees of freedom. The number of observations used by this term.

### Term Fixed?

Indicates whether the term is “fixed” or “random.”

### Denominator Term

Indicates the term used as the denominator in the F-ratio.

### Expected Mean Square

This expression represents the expected value of the corresponding mean square if the design was completely balanced. “S” represents the expected value of the mean square error (experimental variance). The uppercase letters represent either the adjusted sum of squared treatment means if the factor is fixed, or the variance component if the factor is random. The lowercase letter represents the number of levels for that factor, and “s” represents the number of replications of the experimental layout.

These EMS expressions are provided to determine the appropriate error term for each factor. The correct error term for a factor is that term whose EMS is identical except for the factor being tested.

The appropriate error term for factor B is the AB interaction. The appropriate error term for AB is S (mean square error). Since there are zero degrees of freedom for S, the terms A and AB cannot be tested.

# Analysis of Variance Table Section

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (Block) | 2 | 10648.67 | 5324.333 | | | |
| B (Treatment) | 3 | 4650.917 | 1550.306 | 1.09 | 0.421359 | 0.177941 |
| AB | 6 | 8507.333 | 1417.889 | | | |
| S | 0 | 0 | | | | |
| Total (Adjusted) | 11 | 23806.92 | | | | |
| Total | 12 | | | | | |

* Term significant at alpha = 0.05

### Source Term

The source of variation. The term in the model.

### DF

The degrees of freedom. The number of observations used by the corresponding model term.

### Sum of Squares

This is the sum of squares for this term. It is usually included in the ANOVA table for completeness, not for direct interpretation.

### Mean Square

An estimate of the variation accounted for by this term. The sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and the mean square of its corresponding error term. This is also called the F-test value.

### Prob Level

The significance level of the above F-ratio. The probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant. Note that if the value is significant at the specified value of alpha, a star is placed to the right of the F-Ratio.

### Power (Alpha=0.05)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error ($\beta$). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual differences among the population means.

The power value calculated here assumes that the population standard deviation is equal to the observed standard deviation and that the differences among the population means are exactly equal to the difference among the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

Generally, you would consider the power of the test when you accept the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If

you accept the null hypothesis with high power, there is not much left to do. At least you know that the means are not different. However, if you accept the null hypothesis with low power, you can take one or more of the following actions:

1. Increase your alpha level. Perhaps you should be testing at alpha = 0.05 instead of alpha = 0.01. Increasing the alpha level will increase the power.
2. Increase your sample size, which will increase the power of your test if you have low power. If you have high power, an increase in sample size will have little effect.
3. Decrease the magnitude of the variance. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

## Friedman's Rank Test Section

**Treatment Ranks Section**

| Treatment | Number Blocks | Median | Mean of Ranks | Sum of Ranks |
|---|---|---|---|---|
| 1 | 3 | 230 | 2 | 6 |
| 2 | 3 | 245 | 3.333333 | 10 |
| 3 | 3 | 252 | 3 | 9 |
| 4 | 3 | 204 | 1.666667 | 5 |

**Friedman Test Section**

| Ties | Friedman (Q) | DF | Prob Level | Concordance (W) |
|---|---|---|---|---|
| Ignored | 3.400000 | 3 | 0.333965 | 0.377778 |
| Correction | 3.400000 | 3 | 0.333965 | 0.377778 |

| | |
|---|---|
| Multiplicity | 0 |

### Treatment

The level of the treatment (fixed factor) whose statistics are reported on this line.

### Number Blocks

The number of levels (categories) of the block variable (random factor).

### Median

The median value of responses at this treatment level.

### Mean of Ranks

The average of the ranks at this treatment level.

### Sum of Ranks

The sum of the ranks at this treatment level.

### Ties

- **Ignored**
  Statistics on this row are not adjusted for ties.
- **Correction**
  Statistics on this row are adjusted for ties.

### Friedman (Q)

The value of Friedman's Q statistic. This statistic is approximately distributed as a Chi-square random variable with degrees of freedom equal to *k-1*, where *k* is the number of treatments. The Chi-square approximation grows closer as the number of blocks is increased.

### DF

The degrees of freedom. The degrees of freedom is equal to *k-1*, where *k* is the number of treatments.

### Prob Level

The significance level of the *Q* statistic. If this value is less than a specified alpha level (often 0.05), the null hypothesis of equal medians is rejected.

### Concordance (W)

The value of Kendall's coefficient of concordance, which measures the agreement between observers of samples. This value ranges between zero and one. A value of one indicates perfect concordance. A value of zero indicates no agreement or independent samples.

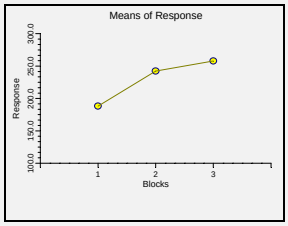
### Multiplicity

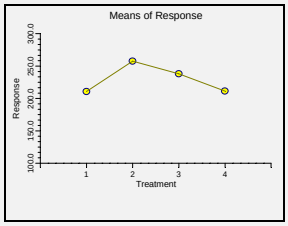
The value of the correction factor for ties: $\sum\left(t^3 - t\right)$.

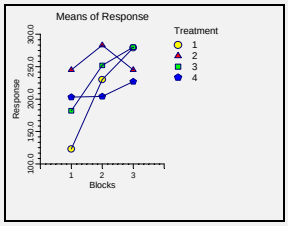
## Means, Effects, and Plots Sections

**Means, Effects, and Plots**

| Term | Count | Mean | Standard Error | Effect |
|---|---|---|---|---|
| All | 12 | 229.4167 | | 229.4167 |
| A: Block | | | | |
| 1 | 4 | 188.25 | 0 | -41.16667 |
| 2 | 4 | 242.25 | 0 | 12.83333 |
| 3 | 4 | 257.75 | 0 | 28.33333 |
| B: Treatment | | | | |
| 1 | 3 | 210.6667 | 21.74005 | -18.75 |
| 2 | 3 | 257.6667 | 21.74005 | 28.25 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |



### Term
The label for this line of the report.

### Count
The number of observations in the mean.

### Mean
The value of the sample mean.

### Standard Error
The standard error of the mean. Note that these standard errors are the square root of the mean square of the error term for this term divided by the count. *These standard errors are not the same as the simple standard errors calculated separately for each group.* The standard errors reported here are those appropriate for testing in multiple comparisons.

### Effect
The additive component that this term contributes to the mean.

### Plot of Means
These plots display the means for each factor and two-way interaction. Note how easily you can see patterns in the plots.

## Multiple-Comparison Sections

**Tukey-Kramer Multiple-Comparison Test**

Response: Response
Term B: Treatment

Alpha=0.050 Error Term=AB DF=6 MSE=1417.889 Critical Value=4.895637

| Group | Count | Mean | Different From Groups |
|---|---|---|---|
| 1 | 3 | 210.6667 | |
| 4 | 3 | 211.3333 | |
| 3 | 3 | 238 | |
| 2 | 3 | 257.6667 | |

These sections present the results of the multiple-comparison procedures selected. These reports all use a uniform format that will be described by considering Tukey-Kramer Multiple-Comparison Test. The reports for the other procedures are similar. For more information on the interpretation of various multiple-comparison procedures, turn to the section by that name in the One-way Analysis of Variance chapter.

### Alpha
The level of significance that you selected.

### Error Term
The term in the ANOVA model that is used as the error term.

### DF
The degrees of freedom of the error term.

### MSE
The value of the mean square error.

### Critical Value
The value of the test statistic that is “just significant” at the given value of alpha. This value depends on which multiple-comparison procedure you are using. It is based on the t-distribution or the studentized range distribution. It is the value of t, F, or q in the corresponding formulas.

### Group
The label for this group.

### Count
The number of observations in the mean.

### Mean
The value of the sample mean.

### Different From Groups
A list of those groups that are significantly different from this group according to this multiple-comparison procedure. All groups not listed are not significantly different from this group.

## Planned-Comparison Section
This section presents the results of any planned comparisons that were selected.

**Planned Comparison: B Linear Trend**

Response: Response
Term B: Treatment

Alpha=0.050 Error Term=AB DF=6 MSE=1417.889

Comparison Value=-3.950387 T-Value=0.1817 Prob>|T|=0.861794 Decision(0.05)=Do Not Reject
Comparison Standard Error=21.74005

| Group | Comparison Coefficient | Count | Mean |
|---|---|---|---|
| 1 | -0.6708204 | 3 | 210.6667 |
| 2 | -0.2236068 | 3 | 257.6667 |
| 3 | 0.2236068 | 3 | 238 |
| 4 | 0.6708204 | 3 | 211.3333 |

### Alpha
The level of significance that you selected.

### Error Term
The term in the ANOVA model that is used as the error term.

### DF
The degrees of freedom of the error term.

### MSE
The value of the mean square error.

### Comparison Value

The value of the comparison. This is formed by multiplying the comparison coefficient by the mean for each group and summing.

### T-Value

The t-test used to test whether the above Comparison Value is significantly different from zero.

$$t_f = \frac{\sum_{i=1}^{k} c_i M_i}{\sqrt{MSE \sum_{i=1}^{k} \frac{c_i^2}{n_i}}}$$

where ***MSE*** is the mean square error, $f$ is the degrees of freedom associated with ***MSE***, $k$ is the number of groups, $c_i$ is the comparison coefficient for the i$^{th}$ group, $M_i$ is the mean of the i$^{th}$ group, and $n_i$ is the sample size of the i$^{th}$ group.

### Prob>|T|

The significance level of the above T-Value. The Comparison is statistically significant if this value is less than the specified alpha.

### Decision(0.05)

The decision based on the specified value of the multiple-comparison alpha.

### Comparison Standard Error

This is the standard error of the estimated comparison value. It is the denominator of the T-Value (above).

### Group

The label for this group.

### Comparison Coefficient

The coefficient (weight) used for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

## Chapter 212

# General Linear Models (GLM)

## Introduction

This procedure performs an analysis of variance or analysis of covariance on up to ten factors using the general linear models approach. The experimental design may include up to two nested terms, making possible various repeated measures and split-plot analyses.

Because the program allows you to control which interactions are included and which are omitted, it can analyze designs with confounding such as Latin squares and fractional factorials.

### Kinds of Research Questions

A large amount of research consists of studying the influence of a set of independent variables on a response (dependent) variable. Many experiments are designed to look at the influence of a single independent variable (factor) while holding other factors constant. These experiments are called single-factor experiments and are analyzed with the one-way analysis of variance (ANOVA). A second type of design considers the impact of one factor across several values of other factors. This experimental design is called the factorial design.

The factorial design is popular among researchers because it not only lets you study the individual effects of several factors in a single experiment, but it also lets you study their interaction. Interaction is present when the response variable fails to behave the same at values of one factor when a second factor is varied. Since factors seldom work independently, the study of their interaction becomes very important.

This procedure will also analyze repeated-measures and split-plot designs. These designs are popular in many disciplines in which experiments are needed that take several measurements on an individual through time. Examples are pre-post type tests administered to various groups of individuals.

Analysis of covariance (ANCOVA) is another design that may be analyzed using this procedure. ANCOVA is useful when you want to improve precision by removing various extraneous sources of variation from your study.

## The Linear Model

We begin with an infinite population of individuals with many measurable characteristics. These individuals are (mentally) separated into two or more treatment populations based on one or more

of these characteristics. A random sample of the individuals in each population is drawn. A treatment is applied to each individual in the sample and an outcome is measured. The data so obtained are analyzed using an analysis of variance table that produces an F-test.

A mathematical model may be formulated that underlies each analysis of variance. This model expresses the response variable as the sum of parameters of the population. For example, a linear mathematical model for a two-factor experiment is

$$Y_{ijk} = m + a_i + b_j + (ab)_{ij} + e_{ijk}$$

where i=1,2,...,I; j=1,2,...,J; and k=1,2,...,K. This model expresses the value of the response variable, *Y*, as the sum of five components:

$m$ the mean.

$a_i$ the contribution of the $i^{th}$ level of a factor A.

$b_j$ the contribution of the $j^{th}$ level of a factor B.

$(ab)_{ij}$ the combined contribution of the $i^{th}$ level of a factor A and the $j^{th}$ level of a factor B.

$e_{ijk}$ the contribution of the $k^{th}$ individual. This is often called the "error."

Note that this model is the sum of various constants. This type of model is called a linear model. It becomes the mathematical basis for our discussion of the analysis of variance. Also note that this serves only as an example. Many linear models could be formulated for the two-factor experiment.

## Assumptions

The following assumptions are made when using the F-test.

1. The response variable is continuous.
2. The $e_{ijk}$ follow the normal probability distribution with mean equal to zero.
3. The variances of the $e_{ijk}$ are equal for all values of i, j, and k.
4. The individuals are independent.

## Limitations

There are few limitations when using these tests. Sample sizes may range from a few to several hundred. If your data are discrete with at least five unique values, you can assume that you have met the continuous variable assumption. Perhaps the greatest restriction is that your data comes from a random sample of the population. If you do not have a random sample, the F-test will not work.

When missing cells occur in your design, you must take special care to be sure that appropriate interaction terms are removed from the ANOVA model.

Special restrictions apply when you are running an analysis with nested terms, as in repeated measures designs. First of all, you cannot have covariates with nested terms. Second, although the sample sizes of groups (the "between" factor) may be unequal, all data must be present for each nested factor. For example, if you are running a pre-post design, you must have both pre- and post- scores for each individual. You cannot include individuals that have only one or the other.

# Multiple Comparison Procedures

The multiple comparison procedures are discussed in the One-Way Analysis of Variance chapter.

# Data Structure

The data must be entered in a format that puts the response in one variable and the values of each of the factors in other variables. An example of the data for a randomized-block design is shown next.

**RNDBLOCK dataset**

| Block | Treatment | Response |
|---|---|---|
| 1 | 1 | 123 |
| 2 | 1 | 230 |
| 3 | 1 | 279 |
| 1 | 2 | 245 |
| 2 | 2 | 283 |
| 3 | 2 | 245 |
| 1 | 3 | 182 |
| 2 | 3 | 252 |
| 3 | 3 | 280 |
| 1 | 4 | 203 |
| 2 | 4 | 204 |
| 3 | 4 | 227 |

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

These panels specify the variables used in the analysis and the model.

### Response Variables

#### Response Variable(s)

Specifies the response (dependent) variable to be analyzed. If you specify more than one variable here, a separate analysis is run for each variable.

### Covariate Specification

#### Covariate(s)

One or more covariates may be specified, causing an analysis of covariance (ANCOVA) to be run. Note that you cannot specify covariates if any of your factors are of the nested type.

## Factor Specification

### Factor Variable

At least one factor variable must be specified. This variable's values indicate how the values of the response variable should be categorized. Examples of factor variables are gender, age groups, "yes" or "no" responses, etc. Note that the values in the variable may be either numeric or text. The treatment of text variables is specified for each variable by the *Data Type* option on the database.

### Type

This option specifies whether the factor is fixed, random, or nested.

A **fixed** factor includes all possible levels, like male and female for gender, includes representative values across the possible range of values, like low, medium, and high temperatures, or includes a set of values to which inferences will be limited, like New York, California, and Maryland.

A **random** factor is one in which the chosen levels represent a random sample from the population of values. For example, you might select four classes from the hundreds in your state or you might select ten batches from an industrial process. The key is that a random sample is chosen. In ***NCSS***, a random factor is "crossed" with other random and fixed factors. Two factors are crossed when each level of one includes all levels of the other.

A **nested** factor is a special type of random factor whose levels (values) are not repeated for all combinations of the factors before it. That is, if factor B is nested in factor A, each level of factor A has its own set of values for factor B.

For example, suppose that factor A represents three fourth-grade classrooms of twenty students in a particular state. Further suppose that factor B represents the sixty children in these classrooms. If factors A and B were crossed, then all sixty children would somehow simultaneously be attending all three classrooms. However, if each classroom has a mutually exclusive set of twenty children, we say that children are nested within classrooms or B is nested within A. Notice that nesting occurs when each level of the first factor (the classrooms) contains separate levels of the second factor (the children).

Note that nested factors should be numbered consecutively, just like random and fixed factors. In the preceding example, you would number the children from one to sixty. You cannot have two individuals with the same identification number.

### Comparisons

Comparisons are only valid for fixed factors. This option lets you specify comparisons that you want to run on this factor. A comparison is formulated in terms of the means as follows:

$$C_i = \sum_{j=1}^{J} w_{ij} m_j$$

In this equation, there are J levels in the factor, the means for each level of the factor are denoted $m_i$, and $w_{ij}$ represents a set of J weight values for the $i^{th}$ comparison. The comparison value, $C_i$, is tested using a t-test. Note that if the $w_{ij}$ sum to zero across j, the comparison is called a "contrast" of the means.

Comparisons may be specified by simply listing the weights. For example, suppose a factor has three levels (unique values). Further suppose that the first level represents a control group, the second a treatment at one dose, and the third a treatment at a higher dose. Three comparisons

come to mind: compare each of the treatment groups to the control group and compare the two treatment groups to each other. These three comparisons would be

| | |
|---|---|
| Control vs. Treatment 1 | -1,1,0 |
| Control vs. Treatment 2 | -1,0,1 |
| Treatment 1 vs. Treatment 2 | 0,-1,1 |

You might also be interested in comparing the control group with the average of both treatment groups. The weights for this comparison would be -2,1,1.

When a factor is quantitative, it might be of interest to divide the response pattern into linear, quadratic, cubic, or other components. If the sample sizes are equal and the factor levels are equally spaced, these so-called components of trend may be studied by the use of simple contrasts. For example, suppose a quantitative factor has three levels: 5, 10, and 15. Contrasts to test the linear and quadratic trend components would be

| | |
|---|---|
| Linear trend | -1,0,1 |
| Quadratic trend | 1,-2,1 |

If the sample sizes for the groups are unequal (the design is unbalanced), adjustments must be made for the differing sample sizes.

***NCSS*** will automatically generate some of the more common sets of contrasts, or it will let you specify up to three custom contrasts yourself. The following common sets are designated by this option.

- **None**

  No comparisons are generated.

- **Standard Set**

  This option generates a standard set of contrasts in which the mean of the first level is compared to the average of the rest, the mean of the second group is compared to the average of those remaining, and so on.

  The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -3,1,1,1 | Compare the first-level mean with the average of the rest. |
  | 0,-2,1,1 | Compare the second-level mean with the average of the rest. |
  | 0,0,-1,1 | Compare the third-level mean with the fourth-level mean. |

- **Polynomial**

  This option generates a set of orthogonal contrasts that allow you to test various trend components from linear up to sixth order. These contrasts are appropriate even if the levels are unequally spaced or the group sample sizes are unequal. Of course, these contrasts are only appropriate for data that are at least ordinal. Usually, you would augment the analysis of this type of data with a multiple regression analysis.

  The following example displays the type of contrasts generated by this option. Suppose there are four equally spaced levels in the factor and each group has two observations. The contrasts generated by this option are (scaled to whole numbers):

  | | |
  |---|---|
  | -3,-1,1,3 | Linear component. |

| | |
|---|---|
| 1,-1,-1,1 | Quadratic component. |
| -1,3,-3,1 | Cubic component. |

- **Linear Trend**

  This option generates a set of orthogonal contrasts and retains only the linear component. This contrast is appropriate even if the levels are unequally spaced and the group sample sizes are unequal. See Orthogonal Polynomials above for more detail.

- **Linear-Quadratic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first two (the linear and quadratic) are reported.

- **Linear-Cubic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first three are reported.

- **Linear-Quartic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first four are reported.

- **Each with First**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the first level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,1,0,0 | Compare the first- and second-level means. |
  | -1,0,1,0 | Compare the first- and third-level means. |
  | -1,0,0,1 | Compare the first- and fourth-level means. |

- **Each with Last**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the last level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -1,0,0,1 | Compare the first- and fourth-level means. |
  | 0,-1,0,1 | Compare the second- and fourth-level means. |
  | 0,0,-1,1 | Compare the third- and fourth-level means. |

- **Custom**

  This option indicates that the contrasts listed in the corresponding three boxes of the Comparison panel should be used.

## Model Specification

This section specifies the experimental design model.

### Which Model Terms

A design in which main effect and interaction terms are included is called a saturated model. Often, it is useful to omit various interaction terms from the model. This option lets you specify which interactions to keep very easily. If the selection provided here is not flexible enough for your needs, you can specify *custom* here and enter the model directly.

The options included here are as follows.

- **Full Model**

  The complete, saturated model is analyzed. This option requires that you have no missing cells, although you can have an unbalanced design. Hence, you cannot use this option with Latin square or fractional factorial designs.

- **Up to 1-Way**

  A main-effects only model is run. All interactions are omitted.

- **Up to 2-Way**

  All main-effects and two-way interactions are included in the model.

- **Up to 3-Way**

  All main-effects, two-way, and three-way interactions are included in the model.

- **Up to 4-Way**

  All main-effects, two-way, three-way, and four-way interactions are included in the model.

- **Custom**

  This option indicates that you want the Custom Model (given in the next box) to be used.

- **Write Model in 'Custom Model' Field**

  When this option is checked, no analysis is performed when the procedure is run. Instead, a copy of the full model is stored in the Custom Model box. You can then delete selected terms from the model without having to enter all the terms you want to keep.

### Custom Model

When Custom Model (see Which Model Terms above) is selected, the model itself is entered here. If all main effects and interactions are desired, you can enter the word "ALL" here. For complicated designs, it is usually easier to check the next option, Write Model in 'Custom Model' Field, and run the procedure. The appropriate model will be generated and placed in this box. You can then edit it as you desire.

The model is entered using letters separated by the plus sign. For example, a three-factor factorial in which only two-way interactions are needed would be entered as follows:

A+B+AB+C+AC+BC.

A simple repeated-measures design would look like this:

A+B(A)+C+AC+BC(A).

# Custom Comparisons Tab

This panel is used when the Comparison option of one or more factors is set to Custom. The Custom option means that the contrast coefficients are to be entered by the user. The boxes on this panel contain the user-supplied contrast coefficients. The first row is for factor one, the second row for factor two, and so on.

## Custom Comparisons

The following options are only used if Comparisons is set to 'Custom' on the Variables tab.

### Custom (1-3)

This option lets you write a user-specified comparison by specifying the weights of that comparison. Note that there are no numerical restrictions on these coefficients. They do not even have to sum to zero. However, this is recommended. If the coefficients do sum to zero, the comparison is called a contrast. The significance tests anticipate that only one or two of these comparisons are to be run. If you run several, you should make some type of Bonferroni adjustment to your alpha value.

When you put in your own contrasts, you must be careful that you specify the appropriate number of weights. For example, if the factor has four levels, four weights must be specified, separated by commas. Extra weights are ignored. If too few weights are specified, the missing weights are set to zero.

These comparison coefficients designate weighted averages of the level-means that are to be statistically tested. The null hypothesis is that the weighted average is zero. The alternative hypothesis is that the weighted average is nonzero. The weights (comparison coefficients) are specified here.

As an example, suppose you want to compare the average of the first two levels with the average of the last two levels in a six-level factor. You would enter “-1,-1,0,0,1,1.”

As a second example, suppose you want to compare the average of the first two levels with the average of the last three levels in a six-level factor. The contrast would be

-3,-3,0,2,2,2.

Note that in each case, we have used weights that sum to zero. This is why we could not use ones in the second example.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Reports

### EMS Report ... Means Report

Specify whether to display the indicated reports.

## Select Plots

### Means Plot(s)

Specify whether to display the indicated plots.

## Report Options

### Test Alpha

The value of alpha for the statistical tests and power analysis. Usually, this number will range from 0.10 to 0.001. A common choice for alpha is 0.05, but this value is a legacy from the age before computers when only printed tables were available. You should determine a value appropriate for your particular study.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, while the double precision will display thirteen-place accuracy.

### Variable Names

Indicate whether to display the variable names or the variable labels.

### Value Labels

Indicate whether to display the data values or their labels.

## Multiple Comparison Tests

### Bonferroni Test (All-Pairs) ... Tukey-Kramer Confidence Intervals

These options specify which MC tests and confidence intervals to display.

### Tests for Two-Factor Interactions

This option specifies whether multiple comparison tests are generated for two-factor interaction terms. When checked, the means of two-factor interactions will be tested by each active multiple comparison test. The multiple comparison test will treat the means as if they came from a single

factor. For example, suppose factor A as two levels and factor B has three levels. The AB interaction would then have six levels. The active multiple comparison tests would be run on these six means.

Care must be used when interpreting multiple comparison tests on interaction means. Remember that the these means contain not only the effects of the interaction, but also the main effects of the two factors. Hence these means contain the combined effects of factor A, factor B, and the AB interaction. You cannot interpret the results as representing only the AB interaction.

## Multiple Comparison Tests – Options

### MC Alpha

Specifies the alpha value used by the multiple-comparison tests.

### MC Decimals

Specify how many decimals to display in the multiple comparison sections.

# Means Plot Tab

These options specify the plots of group means.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Y Scaling

Specify the method for calculating the minimum and maximum along the vertical axis. *Separately* means that each plot is scaled independently. *Uniform* means that all plots use the overall minimum and maximum of the data. This option is ignored if a minimum or maximum is specified.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Lines

Click this box to connect the points for a particular factor. This makes it easier to spot patterns in the means.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, *{G}* is replaced by the appropriate factor name.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

## Symbols Tab

These options specify the symbols used in the plots of group means.

### Plotting Symbols

#### Group (1-15)

The symbols used to represent the levels of a factor on the means plots. Group 1 represents the first level, Group 2 represents the second level, and so on.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

#### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running a GLM ANOVA

This section presents an example of how to run an analysis of the data presented in Table 212.1. These data are contained in the RNDBLOCK database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the General Linear Models (GLM) window.

### 1 Open the RNDBLOCK dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **RNDBLOCK.s0**.
- Click **Open**.

### 2 Open the General Linear Models (GLM) window.

- On the menus, select **Analysis**, then **Analysis of Variance (ANOVA)**, then **General Linear Models (GLM)**. The General Linear Models (GLM) procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the General Linear Models (GLM) window, select the **Variables tab**.
- Double-click in the **Response Variables** box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**.
- Double-click in the **Factor 1 Variable** box. This will bring up the variable selection window.
- Select **Block** from the list of variables and then click **Ok**.
- Select **Random** in the **Type** box for Factor 1.
- Double-click in the **Factor 2 Variable** box. This will bring up the variable selection window.
- Select **Treatment** from the list of variables and then click **Ok**.
- Select **Fixed** in the **Type** box for Factor 2.
- Select **Linear** in the **Comparisons** box for Factor 2.

### 4 Specify the reports.

- On the General Linear Models (GLM) window, select the **Reports tab**.
- Check the **Tukey-Kramer Test** option of the Multiple Comparison Tests.

### 5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

We will now document this output, one section at a time.

# Expected Mean Squares Section

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (Block) | 2 | No | S(AB) | S+bsA |
| B (Treatment) | 3 | Yes | AB | S+sAB+asB |
| AB | 6 | No | S(AB) | S+sAB |
| S(AB) | 0 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

The expected mean square expressions are provided to show the appropriate error term for each factor. The correct error term for a factor is that term that is identical except for the factor being tested.

### Source Term

The source of variation or term in the model.

### DF

The degrees of freedom, which is the number of observations used by this term.

### Term Fixed?

Indicates whether the term is fixed or random.

### Denominator Term

Indicates the term used as the denominator in the F-ratio.

### Expected Mean Square

This expression represents the expected value of the corresponding mean square if the design was completely balanced. *S* represents the expected value of the mean square error (sigma). The uppercase letters represent either the adjusted sum of squared treatment means if the factor is fixed, or the variance component if the factor is random. The lowercase letter represents the number of levels for that factor, and *s* represents the number of replications of the experimental layout.

These EMS expressions are provided to determine the appropriate error term for each factor. The correct error term for a factor is that term whose EMS is identical except for the factor being tested.

In this example, the appropriate error term for factor B is the AB interaction. The appropriate error term for AB is S (mean square error). Since there are zero degrees of freedom for S, the terms A and AB cannot be tested.

# Analysis of Variance Table Section

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (Blocks) | 2 | 10648.67 | 5324.333 | | | |
| B (Treatment) | 3 | 4650.917 | 1550.306 | 1.09 | 0.421359 | 0.177941 |
| AB | 6 | 8507.333 | 1417.889 | | | |
| S | 0 | 0 | | | | |
| Total (Adjusted) | 11 | 23806.92 | | | | |
| Total | 12 | | | | | |

* Term significant at alpha = 0.05

### Source Term

The source of variation, which is the term in the model.

### DF

The degrees of freedom, which is the number of observations used by the corresponding model term.

### Sum of Squares

This is the sum of squares for this term. It is usually included in the ANOVA table for completeness, not for direct interpretation.

### Mean Square

An estimate of the variation accounted for by this term; it is the sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and the mean square of its corresponding error term. This is also called the F-test value.

### Prob Level

The significance level of the above F-ratio, or the probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant. Note that if the value is significant at the specified value of alpha, a star is placed to the right of the F-Ratio.

### Power (Alpha=0.05)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error (β). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual differences among the population means.

The power value calculated here assumes that the population standard deviation is equal to the observed standard deviation and that the differences among the population means are exactly equal to the differences among the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

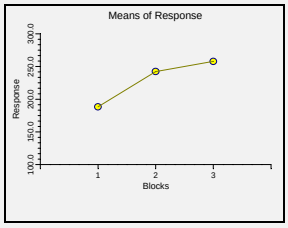
Generally, you would consider the power of the test when you accept the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If you accept the null hypothesis with high power, there is not much left to do. At least you know that the means are not different. However, if you accept the null hypothesis with low power, you can take one or more of the following actions:

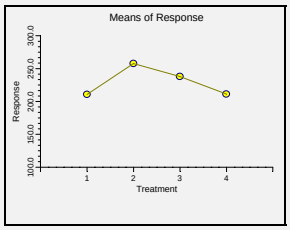
1. Increase your alpha level. Perhaps you should be testing at alpha = 0.05 instead of alpha = 0.01. Increasing the alpha level will increase the power.
2. Increasing your sample size will increase the power of your test if you have low power. If you have high power, an increase in sample size will have little effect.
3. Decrease the magnitude of the variance. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

## Means and Standard Errors Section, Plots Sections

**Means and Standard Errors , and Plots Sections**

| Term | Count | Mean | Standard Error |
|---|---|---|---|
| All | 12 | 229.4167 | |
| A: Blocks | | | |
| 1 | 4 | 188.25 | 0 |
| 2 | 4 | 242.25 | 0 |
| 3 | 4 | 257.75 | 0 |
| B: Treatment | | | |
| 1 | 3 | 210.6667 | 21.74005 |
| 2 | 3 | 257.6667 | 21.74005 |
| . | . | . | . |
| . | . | . | . |




### Term

The label for this line of the report.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Standard Error

The standard error of the mean. Note that these standard errors are the square root of the mean square of the error term for this term divided by the count. These standard errors are not the same as the simple standard errors calculated separately for each group. The standard errors reported here are those appropriate for testing multiple comparisons.

Note that the standard errors for the means of Block are zero since there is no error term for this factor. This may be seen by looking at the Expected Mean Squares Report above.

### Plot of Means

These plots display the means for each factor and two-way interactions. Note how easily you can see patterns in the plots.

## Multiple-Comparison Sections

**Tukey-Kramer Multiple-Comparison Test**

Response: Response
Term B: Treatment

Alpha=0.050 Error Term=AB DF=6 MSE=1417.889 Critical Value=4.8956

| Group | Count | Mean | Different From Groups |
|---|---|---|---|
| 1 | 3 | 210.6667 | |
| 4 | 3 | 211.3333 | |
| 3 | 3 | 238 | |
| 2 | 3 | 257.6667 | |

These sections present the results of the multiple-comparison procedures selected. These reports all use a uniform format that will be described by considering Tukey-Kramer Multiple-Comparison Test. The reports for the other procedures are similar. For more information on the interpretation of the various multiple-comparison procedures, turn to the section by that name in the One-Way ANOVA chapter.

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom for the error term.

### MSE

The value of the mean square error.

### Critical Value

The value of the test statistic that is “just significant” at the given value of alpha. This value depends on which multiple-comparison procedure you are using. It is based on the t-distribution or the studentized range distribution. It is the value of t, F, or q in the corresponding formulas.

### Group

The label for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Different from Groups

A list of those groups that are significantly different from this group according to this multiple-comparison procedure. All groups not listed are not significantly different from this group.

## Planned-Comparison Section

This section presents the results of any planned comparisons that were selected.

**Planned Comparison: B Linear Trend**

Response: Response
Term B: Treatment

Alpha=0.050 Error Term=AB DF=6 MSE=1417.889

Comparison Value=-3.950387 T-Value=0.1817 Prob>|T|=0.861794 Decision(0.05)=Do Not Reject
Comparison Standard Error=21.74005

| Group | Comparison Coefficient | Count | Mean |
|---|---|---|---|
| 1 | -0.6708204 | 3 | 210.6667 |
| 2 | -0.2236068 | 3 | 257.6667 |
| 3 | 0.2236068 | 3 | 238 |
| 4 | 0.6708204 | 3 | 211.3333 |

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom of the error term.

### MSE

The value of the mean square error.

### Comparison Value

The value of the comparison. This is formed by multiplying the Comparison Coefficient times the Mean for each group and summing.

### T-Value

The t-test used to test whether the above Comparison Value is significantly different from zero.

$$t_f = \frac{\sum_{i=1}^{k} c_i M_i}{\sqrt{MSE \sum_{i=1}^{k} \frac{c_i^2}{n_i}}}$$

where ***MSE*** is the mean square error, ***f*** is the degrees of freedom associated with ***MSE***, ***k*** is the number of groups, $c_i$ is the comparison coefficient for the $i^{th}$ group, $M_i$ is the mean of the $i^{th}$ group, and $n_i$ is the sample size of the $i^{th}$ group.

### Prob>|T|

The significance level of the above T-Value. The Comparison is statistically significant if this value is less than the specified alpha.

### Decision(0.05)

The decision based on the specified value of the multiple comparison alpha.

### Comparison Standard Error

This is the standard error of the estimated comparison value. It is the denominator of the T-Value (above).

### Group

The label for this group.

### Comparison Coefficient

The coefficient (weight) used for this group. Note that for our example, the weights are appropriate for the linear-trend component of a set of orthogonal polynomials.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

# GLM ANOVA Checklist

This checklist, prepared by a professional statistician, is a flowchart of the steps you should complete to conduct a valid analysis. Since this topic is vast, this flowchart will give only a brief summary. You should consult appropriate statistical books in your field for further details. We recommend Winer (1990) and Keppel (1991) as good books to use, but there are many others available that are equally useful.

## Step 1 – Data Preparation

### Introduction

This step involves scanning your data for anomalies, keypunch errors, typos, and so on. You would be surprised how often we hear of people completing an analysis, only to find that they had mistakenly selected the wrong variables.

### Sample Size

The sample size (number of nonmissing rows) has a lot of ramifications. The analysis of variance was originally developed under the assumption that the sample sizes of each treatment combination are equal. In practice this seldom happens, but the closer you can get to equal sample sizes the better.

### Missing Values

The number and pattern of missing values are always issues to consider. Usually, we assume that missing values occur at random throughout your data. If this is not true, your results will be biased since a particular segment of the population is underrepresented.

If you have missing values, it will be important to identify the degree of unbalance in your design. You should also check to see if there are any missing cells. If there are, you cannot run a full model. You will have to assume some interactions are zero and remove them from the ANOVA model.

### Type of Data

The mathematical basis of the F-test assumes that the data are continuous. Because of the rounding that occurs when data are recorded, all data are technically discrete. The validity of assuming the continuity of the data then comes down to determining when we have too much rounding. For example, most statisticians would not worry about human-age data that was rounded to the nearest year. However, if these data were rounded to the nearest ten years or further to only three groups (young, adolescent, and adult), most statisticians question the validity of the probability statements. Some studies have shown that the F-test is reasonably accurate when the data have only five possible values (most would call this discrete data). If your data contain less than five unique values, any probability statements made are tenuous.

Also, you should double-check to ensure that you are going to use the appropriate design. Our experience is that many researchers use a factorial design when they should be using a repeated measures design. Consider again the examples of each type of design and make sure you are using the correct one.

### Outliers

Generally, outliers cause distortion in most popular statistical tests. You must scan your data for outliers (the box plot is an excellent tool for doing this). If you have outliers, you have to decide if they are one-time occurrences or if they would occur in another sample. If they are one-time occurrences, you can remove them and proceed. If you know they represent a certain segment of the population, you have to decide between biasing your results (by removing them) or leaving them in and invalidating the normality assumption.

## Step 2 – Setup and Run the GLM ANOVA Panel

### Introduction

Now comes the fun part: running the program. ***NCSS*** is designed to be simple to operate, but it can still seem complicated. When you go to run a procedure such as this for the first time, take a few minutes to read through the chapter again and familiarize yourself with the issues involved.

### Enter Variables

The templates are set with ready-to-run defaults. About all you have to do is select the appropriate variables (columns of data).

### Select All Plots

As a rule, you should select the means plots. They add a great deal to your ability to interpret the data.

### Specify Alpha

Most beginners at statistics forget this important step and let the alpha value default to the standard 0.05. You should consciously decide what value of alpha is appropriate for your study. The 0.05 default came about when people had to rely on printed probability tables and there were only two values available: 0.05 or 0.01. Now you can set the value to whatever is appropriate.

A special note on setting the Multiple Comparison alpha. You will often want to reset this value to 0.10 so that the individual tests are made at a more reasonable significance level.

## Step 3 – Check Assumptions

### Introduction

Testing the assumptions of normality and equal variance is often difficult in a multi-way analysis of variance. We suggest that you make several passes through your data using our one-way ANOVA program, studying each factor separately. We suggest this because the one-way ANOVA program displays extensive diagnostic information for checking equal variance and normality. Although this method does not account for the interactions among the factors, it is often the best you can do to assess the validity of your assumptions.

Sometimes, the ANOVA model can be recoded so that you can run it through our regression program. When this is possible, you can analyze the residuals to assess normality and equal variance.

### Random Sample

These statistical procedures were designed with the assumption that the sample population was selected randomly. The validity of this assumption depends on the method used to select the sample. If you have not used valid sampling techniques, the F-test will not work.

### Check Descriptive Statistics

You should check the Means and Standard Errors Section first to determine if the Counts and the Means are reasonable. If you have selected the wrong variable, these values will alert you.

## Step 4 – Interpret Findings

### Introduction

You are now ready to conduct your tests. The basic plan of attack for analyzing your output is as follows:

1. Glance through the reports, checking the means, the F-tests, and so forth for obvious problems.
2. Look at the power of the nonsignificant tests. Could the lack of significance be the result of a small sample size?

3. Determine which main effects and interactions are significant.
4. Use care in interpreting a main effect when its interaction with another term is significant.
5. Use planned comparisons, paired comparisons, and plots of means to view the experimental results and discuss what they reveal.

# Examples of Various Experimental Designs

We will now present examples of how to run various popular types of experimental designs.

## Randomized-Block Design

The randomized-block design is a very popular experimental design. The focus of the analysis is on a set of two or more treatments. A blocking variable is used to account for extraneous factors. Each block receives all treatments. These treatments are randomly assigned within the block.

The data in the RNDBLOCK dataset show how to enter the data for this type of design. You should designate the block term as *random* and the treatment term as *fixed*. Set the Which Model Terms option to Up to 1-Way (removing the interaction term). In a typical randomized-block design, the interaction term becomes the error term, so it does not have to be fit separately. Doing this will reduce the amount of time needed to complete the calculations.

## Single-Factor Repeated-Measures Design

The single-factor repeated-measures design is similar to the randomized-block design. In this design the individuals (analogous to the blocks) are measured over time. Unlike the randomized-block design, however, the treatments are not applied in random order. Instead, the treatments are always applied in the same order. For example, you might conduct a pre-test, apply some treatment to the individuals, and conduct a post-test. You cannot apply the post-test first.

The data in the RNDBLOCK dataset show how to enter the data for this type of design if you think of blocks as the individuals and treatments as time of measurement.

It turns out that even though the randomization method is different, the analysis of this design is identical to that described above for the randomized-block design. The individuals become the blocks. This variable is designated *random*. The repeated-measures variable (the variable representing time) becomes the treatment. This variable is designated as *fixed*. Set the Which Model Terms option (Model Tab) to Up to 1-Way to omit the interaction term.

## Latin-Square Design

Fractional-rep designs are known for their ability to provide insight about several factors with a minimum number of observations. This efficiency comes from an experimental setup that ignores many interaction terms. The Latin square is one such design. It may be analyzed with ***NCSS***. The following table shows a set of Latin-square data from page 313 of Snedecor and Cochran (1972).

**Latin-Square Data from Snedecor and Cochran**

| Row | Column 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **1** | B:257 | E:230 | A:279 | C:287 | D:202 |
| **2** | D:245 | A:283 | E:245 | B:280 | C:260 |
| **3** | E:182 | B:252 | C:280 | D:246 | A:250 |
| **4** | A:203 | C:204 | D:227 | E:193 | B:259 |
| **5** | C:231 | D:271 | B:266 | A:334 | E:338 |

The following table shows the data as it would be entered for analysis in ***NCSS***. The *Custom Model* statement "A+B+C" would be used since many of the interactions cannot be estimated. The factors would be designated as fixed or random depending on the experimental situation.

**LATINSQR dataset**

| Rows | Columns | Letters | Yield |
|---|---|---|---|
| 1 | 1 | B | 257 |
| 1 | 2 | E | 230 |
| 1 | 3 | A | 279 |
| 1 | 4 | C | 287 |
| 1 | 5 | D | 202 |
| 2 | 1 | D | 245 |
| 2 | 2 | A | 283 |
| 2 | 3 | E | 245 |
| 2 | 4 | B | 280 |
| 2 | 5 | C | 260 |
| 3 | 1 | E | 182 |
| 3 | 2 | B | 252 |
| 3 | 3 | C | 280 |
| 3 | 4 | D | 246 |
| 3 | 5 | A | 250 |
| 4 | 1 | A | 203 |
| 4 | 2 | C | 204 |
| 4 | 3 | D | 227 |
| 4 | 4 | E | 193 |
| 4 | 5 | B | 259 |
| 5 | 1 | C | 231 |
| 5 | 2 | D | 271 |
| 5 | 3 | B | 266 |
| 5 | 4 | A | 334 |
| 5 | 5 | E | 338 |

## Repeated-Measures Design

A Repeated Measures ANOVA is a particular type of three-factor design that uses two error terms. In this design, treatments are applied to experimental units of different sizes. For example, in an educational study, one treatment might be applied to whole classrooms. A second treatment might consist of the students' responses to a pre-test and a post-test. Such a design employs two error terms. One error term is the between-classes error for testing the first factor. The other error term is the within-student error for testing the second factor.

This procedure analyzes data from an experimental design represented by the following mathematical model:

$$Y_{ijkl} = \mu + A_i + S_{ij} + B_k + AB_{ik} + e_{ijkl}$$

In this model, A is the between-group treatment, $S_{ij}$ is the between-group error, B is the within-subject treatment, and $e_{ijkl}$ is the within-subject error.

This is a specialized technique with strict assumptions. An advanced statistical text dealing with the topic should be consulted before the technique is employed.

The data below illustrate how the data should be set up. An experiment was conducted to study the effects of exercise on heart rate. The subjects were randomly divided into three groups of six. The first group did not have a regular exercise plan. The second group exercised once a week. The third group exercised daily. Each subject's heart rate was recorded when the experiment began and again at the end of ten weeks. These data are stored in a database called HEART. You might want to open this database and run the analysis yourself.

**HEART dataset**

| Exercise | Subject | Time | Heart Rate |
|---|---|---|---|
| None | 1 | 0 | 87 |
| None | 1 | 10 | 89 |
| None | 2 | 0 | 67 |
| None | 2 | 10 | 65 |
| None | 3 | 0 | 55 |
| None | 3 | 10 | 58 |
| None | 4 | 0 | 66 |
| None | 4 | 10 | 68 |
| None | 5 | 0 | 88 |
| None | 5 | 10 | 90 |
| None | 6 | 0 | 75 |
| None | 6 | 10 | 73 |
| Weekly | 7 | 0 | 84 |
| Weekly | 7 | 10 | 78 |
| Weekly | 8 | 0 | 78 |
| Weekly | 8 | 10 | 72 |
| Weekly | 9 | 0 | 64 |
| Weekly | 9 | 10 | 53 |
| Weekly | 10 | 0 | 73 |
| Weekly | 10 | 10 | 65 |
| Weekly | 11 | 0 | 84 |
| Weekly | 11 | 10 | 82 |
| Weekly | 12 | 0 | 55 |

**HEART dataset (continued)**

| Exercise | Subject | Time | Heart Rate |
|---|---|---|---|
| Weekly | 12 | 10 | 53 |
| Daily | 13 | 0 | 72 |
| Daily | 13 | 10 | 55 |
| Daily | 14 | 0 | 83 |
| Daily | 14 | 10 | 72 |
| Daily | 15 | 0 | 75 |
| Daily | 15 | 10 | 63 |
| Daily | 16 | 0 | 55 |
| Daily | 16 | 10 | 49 |
| Daily | 17 | 0 | 83 |
| Daily | 17 | 10 | 68 |
| Daily | 18 | 0 | 63 |
| Daily | 18 | 10 | 54 |

To run this analysis, you specify *Heart Rate* as the Response Variable, *Exercise* as Factor 1 (designate it as fixed), *Subject* as Factor 2 (designate it as nested), and *Time* as Factor 3 (designate it as fixed). Select the full model. When the analysis is complete, the following output is displayed.

## Repeated-Measure ANOVA Report for Heart Rate Data

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Exercise | 2 | 331.1667 | 165.5833 | 0.61 | 0.555761 | 0.133134 |
| B(A): Subject | 15 | 4064.833 | 270.9889 | | | |
| C: Time | 1 | 277.7778 | 277.7778 | 50.51 | 0.000004* | 0.999998 |
| AC | 2 | 234.7222 | 117.3611 | 21.34 | 0.000041* | 0.999793 |
| BC(A) | 15 | 82.5 | 5.5 | | | |
| S | 0 | 0 | | | | |
| Total (Adjusted) | 35 | 4991 | | | | |
| Total | 36 | | | | | |

* Term significant at alpha = 0.05

**Means and Effects Section**

| Term | Count | Mean | Standard Error |
|---|---|---|---|
| All | 36 | 69.83334 | |
| A: Exercise | | | |
| Daily | 12 | 66 | 4.752095 |
| None | 12 | 73.41666 | 4.752095 |
| Weekly | 12 | 70.08334 | 4.752095 |
| C: Time | | | |
| 0 | 18 | 72.61111 | .5527708 |
| 10 | 18 | 67.05556 | .5527708 |
| AC: Exercise,Time | | | |
| Daily,0 | 6 | 71.83334 | .9574271 |
| Daily,10 | 6 | 60.16667 | .9574271 |
| None,0 | 6 | 73 | .9574271 |
| None,10 | 6 | 73.83334 | .9574271 |
| Weekly,0 | 6 | 73 | .9574271 |
| Weekly,10 | 6 | 67.16666 | .9574271 |

## Analysis of Covariance

The analysis of covariance uses features from both analysis of variance and multiple regression. The usual one-way classification model in analysis of variance is

$$Y_{ij} = \mu_i + e_{1ij}$$

where $Y_{ij}$ is the $j^{th}$ observation in the $i^{th}$ group, $\mu_i$ represents the true mean of the $i^{th}$ group, and $e_{ij}$ are the residuals or errors in the above model (usually assumed to be normally distributed). Suppose you have measured a second variable with values $X_{ij}$ that is linearly related to Y. Further suppose that the slope of the relationship between Y and X is constant from group to group. You could then write the analysis of covariance model

$$Y_{ij} = \mu_i + ß(X_{ij}-X..) + e_{2ij}$$

where X.. represents the overall mean of X. If X and Y are closely related, you would expect that the errors, $e_{2ij}$, would be much smaller than the errors, $e_{1ij}$, giving you more precise results.

The analysis of covariance is useful for many reasons, but it does have the (highly) restrictive assumption that the slope is constant over all the groups. This assumption is often violated, which limits the technique's usefulness. You will want to study more about this technique in statistical texts before you use it.

Running an analysis of covariance is easy in ***NCSS***. You fill out the procedure template as usual for an ANOVA. To change your ANOVA into an ANCOVA, you simply specify one or more covariates. We will now take you through an extended example showing how to run an Ancova as well as how to test the assumption of equal slopes. The following data give the home state, age, and IQ of thirty teenagers. The variables X1-X4 are for use in testing the Ancova assumption of equal slopes and they will be explained later.

Suppose we wish to test for differences in IQ among the three states while controlling for age (the covariate). These data are contained in the ANCOVA database. You should open this database now if you want to follow along.

**ANCOVA dataset**

| State | Age | IQ | X1 | X2 | X3 | X4 |
|---|---|---|---|---|---|---|
| Iowa | 12 | 100 | -1 | -1 | -12 | -12 |
| Iowa | 13 | 102 | -1 | -1 | -13 | -13 |
| Iowa | 12 | 97 | -1 | -1 | -12 | -12 |
| Iowa | 14 | 96 | -1 | -1 | -14 | -14 |
| Iowa | 15 | 105 | -1 | -1 | -15 | -15 |
| Iowa | 18 | 106 | -1 | -1 | -18 | -18 |
| Iowa | 12 | 105 | -1 | -1 | -12 | -12 |
| Iowa | 14 | 103 | -1 | -1 | -14 | -14 |
| Iowa | 12 | 99 | -1 | -1 | -12 | -12 |
| Iowa | 10 | 98 | -1 | -1 | -10 | -10 |
| Utah | 14 | 104 | 0 | 2 | 0 | 28 |
| Utah | 11 | 105 | 0 | 2 | 0 | 22 |
| Utah | 12 | 106 | 0 | 2 | 0 | 24 |
| Utah | 15 | 103 | 0 | 2 | 0 | 30 |
| Utah | 17 | 102 | 0 | 2 | 0 | 34 |
| Utah | 18 | 99 | 0 | 2 | 0 | 36 |
| Utah | 19 | 107 | 0 | 2 | 0 | 38 |

**ANCOVA dataset (continued)**

| State | Age | IQ | X1 | X2 | X3 | X4 |
|---|---|---|---|---|---|---|
| Utah | 16 | 105 | 0 | 2 | 0 | 32 |
| Utah | 15 | 103 | 0 | 2 | 0 | 30 |
| Utah | 14 | 103 | 0 | 2 | 0 | 28 |
| Texas | 15 | 105 | 1 | -1 | 15 | -15 |
| Texas | 16 | 106 | 1 | -1 | 16 | -16 |
| Texas | 12 | 103 | 1 | -1 | 12 | -12 |
| Texas | 13 | 99 | 1 | -1 | 13 | -13 |
| Texas | 14 | 93 | 1 | -1 | 14 | -14 |
| Texas | 11 | 104 | 1 | -1 | 11 | -11 |
| Texas | 18 | 103 | 1 | -1 | 18 | -18 |
| Texas | 19 | 100 | 1 | -1 | 19 | -19 |
| Texas | 18 | 101 | 1 | -1 | 18 | -18 |
| Texas | 16 | 104 | 1 | -1 | 16 | -16 |

We begin by loading the ANCOVA database and the *GLM ANOVA* options panel. We specify *IQ* as the *Response Variable*, *Age* as the *Covariate*, and *State* as *Factor 1*. We run the procedure and after a few seconds, the analysis of covariance table is displayed.

## Analysis of Covariance Report

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| X(Age) | 1 | 5.239314 | 5.239314 | 0.47 | 0.497230 | 0.101741 |
| A: State | 2 | 28.38448 | 14.19224 | 1.28 | 0.293886 | 0.253298 |
| S | 26 | 287.3607 | 11.05233 | | | |
| Total (Adjusted) | 29 | 328.8 | | | | |
| Total | 30 | | | | | |

Notice that now, in addition to the test for factor A, we also have a test for the covariate. This test, the one along the line labeled "X(AGE)," tests the significance of the covariate. If it is not significant (as is the case in this example), analysis of covariance should not be used. However, if it is significant, you may proceed to the next F-test, the one dealing with factor A (STATE). This is the test that is usually desired in the analysis of covariance. It tests whether the adjusted means of the three states are different. The means are adjusted as if all three states had the same age. That is, the means for each state are adjusted to the average value of age. These adjusted means are shown in the *Means and Effects* report. If you run the analysis without the covariate, you'll notice that these means are different.

Since the covariate (AGE) is not significant, you should stop here. However, for the sake of instruction, we will assume that the covariate is significant and proceed to test whether the slopes between IQ and AGE are the same in the three states. The following steps will lead you through this test:

1. Construct a new contrast variable for each degree of freedom of the factor. In our current example, the three levels (states) of factor A yield two degrees of freedom, so we must create two contrast variables. These are shown in Table 212.5 as X1 and X2.
2. Multiply each of these new variables by the covariate variable. In our example, X3=(X1)(Age) and X4=(X2)(Age).

3. Run another Ancova, using the same setup as before except now you fit the three covariates AGE, X3, and X4. Call these the Model2 results, and call the previous results with just the single covariate the Model1 results.

## Second Analysis of Covariance Report

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob>F | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| X(Age) | 1 | 9.740934 | 9.740934 | 0.94 | 0.341886 | 0.153727 |
| X(X3) | 1 | 22.27164 | 22.27164 | 2.15 | 0.155572 | 0.290793 |
| X(X4) | 1 | 21.07455 | 21.07455 | 2.03 | 0.166672 | 0.277862 |
| A: State | 2 | 46.57466 | 23.28733 | 2.25 | 0.127408 | 0.412350 |
| S | 24 | 248.6402 | 10.36001 | | | |
| Total (Adjusted) | 29 | 328.8 | | | | |
| Total | 30 | | | | | |

4. Finally, create the F-test for equality of slopes as follows. The formula is

$$F_{k,m} = \frac{(SSE_1 - SSE_2)/k}{MSE_2}$$

where k is the degrees of freedom of the factor (in our example, this is 2), m is the degrees of freedom of the mean square for error in model2, $SSE_1$ and $SSE_2$ are the sums of squares error for model1 and model2, and $MSE_2$ is the mean square for error in model2.

The calculations for this example proceed as follows:

$$F_{2,24} = [(287.3607-248.6402)/2]/10.36001 = 1.86875.$$

This F-ratio would then be compared against a tabulated 0.05 F-value, 3.403, which you could find in a statistics book. Since 1.86875 < 3.403, we would not reject the equality of slopes assumption in this case.

One final note, you should generate a scatter plot, which shows the response variable on the vertical axis, the covariate on the horizontal axis, and uses different symbols for each group. The least squares trend line can also be displayed. This plot will let you visually assess the validity of the assumption of equal slope.

# Hierarchical-Classification Designs

Snedecor and Cochran (1967), page 286, present an example of a hierarchical-classification design. In this example, four plants were selected at random, and three leaves were randomly selected from each plant. Two samples were taken from each leaf, and the amount of calcium in the sample was recorded. The data are displayed below. The data are stored in a database called PLANT.

**PLANT dataset**

| Row | Plant | Leaf | Calcium |
|---|---|---|---|
| 1 | 1 | 1 | 3.28 |
| 2 | 1 | 1 | 3.09 |
| 3 | 1 | 2 | 3.52 |
| 4 | 1 | 2 | 3.48 |
| 5 | 1 | 3 | 2.88 |
| 6 | 1 | 3 | 2.80 |
| 7 | 2 | 4 | 2.46 |
| 8 | 2 | 4 | 2.44 |
| 9 | 2 | 5 | 1.87 |
| 10 | 2 | 5 | 1.92 |
| 11 | 2 | 6 | 2.19 |
| 12 | 2 | 6 | 2.19 |
| 13 | 3 | 7 | 2.77 |
| 14 | 3 | 7 | 2.66 |
| 15 | 3 | 8 | 3.74 |
| 16 | 3 | 8 | 3.44 |
| 17 | 3 | 9 | 2.55 |
| 18 | 3 | 9 | 2.55 |
| 19 | 4 | 10 | 3.78 |
| 20 | 4 | 10 | 3.87 |
| 21 | 4 | 11 | 4.07 |
| 22 | 4 | 11 | 4.12 |
| 23 | 4 | 12 | 3.31 |
| 24 | 4 | 12 | 3.31 |

To run this analysis, you specify *Calcium* as the Response Variable, *Plant* as Factor 1 (designate it as random), and *Leaf* as Factor 2 (designate it as nested). Select the *full model*. When the analysis is complete, the following output is displayed.

## Plant Data Example

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (Plant) | 3 | No | B(A) | S+sB+bsA |
| B(A) | 8 | No | S(AB) | S+sB |
| S(AB) | 12 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (Plant) | 3 | 7.560346 | 2.520115 | 7.67 | .009725* | |
| B(A) | 8 | 2.6302 | .328775 | 49.41 | .000000* | |
| S(AB) | 12 | .07985 | 6.654167E-03 | | | |
| Total (Adjusted) | 23 | 10.2704 | | | | |
| Total | 24 | | | | | |

# Chapter 213

# Analysis of Two-Level Designs

## Introduction

Several analysis programs are provided for the analysis of designed experiments. The GLM-ANOVA and the Multiple Regression programs are often used. This chapter describes a program to analyze very particular designs: two-level factorials (with an optional blocking variable) in which the number of rows is a power of two (4, 8, 16, 32, 64, 128, etc.) and there are no missing values.

Given that your data meet these restrictions, this program gives you a complete analysis including:

1. Analysis of the design itself.
2. List of confounding and aliasing patterns.
3. Analysis of variance table.
4. Tables of means and effects.
5. Probability plots of residuals and effects.
6. Two-way and cube plots of means and differences.

## Procedure Options

This section describes the options available in this procedure.

### Variables Tab

These options specify the variables that will be used in the analysis. They also specify the type of analysis that will be performed.

#### Response Variable

**Response Variable**

Specifies the response (dependent) variable to be analyzed.

## Block Variable

### Block Variable

An optional variable containing the levels of the blocking factor. Note that block sizes must be a power of two.

## Factor Specification

### Factor Variables

At least two factor (categorical) variables must be specified. Each factor consists of a variable that contains a column of two unique values (two levels). The values may be text or numeric.

## Error Estimation Options

### Pooled Terms

Often, two-level designs do not provide a direct estimate of the mean square error (MSE). The F-tests in the analysis of variance require an estimate of the MSE, so this option lets you specify one.

This option provides a list of term numbers (separated by commas) that represent the terms that should be pooled (averaged) to form the estimated MSE. These should be determined from the probability plot of the effects and from the Sorted Means and Effects report. This is a list of the terms whose effect is small in absolute value.

This list is optional and may be left blank, in which case it will be ignored. Note that this list is also ignored when the Estimated MSE option is non-zero.

### Estimated MSE

Often, two-level designs do not provide a direct estimate of the mean square error (MSE). The F-tests in the analysis of variance require an estimate of the MSE, so this option lets you specify one.

This option allows the direct specification of an MSE value. This value overrides the Pooled Terms option when it is nonzero. The degrees of freedom associated with the MSE are set to 99. This MSE value should be determined from the analysis of variance of previous experiments.

This value is optional and may be left at zero, in which case it will be ignored.

### Error DF

Enter a value for the error degrees of freedom. This value is only used when the 'Estimated MSE' is non-zero.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Additional Reports

### Show Two-Way Tables

Display the Two-Way Tables of Means and Effects. These reports are useful in analyzing the two-way interactions.

### Show Three-Way Tables

Display the Three-Way Tables of Means and Effects. These reports are useful in analyzing the three-way interactions.

## Select Plots

### Show Probability Plots

Specify whether to display probability plots of the residuals and effects.

### Show Means Plots

Specify whether to display plots of the means for all factors and two-way interactions.

## Report Options

### Confounding Size

Specify what order interactions are included in the report showing the confounding and aliasing patterns in the design.

### Alpha Level

The value of alpha for the statistical tests. Usually, this number will range from 0.1 to 0.001. A common choice for alpha is 0.05, but this value is a legacy from the age before computers when only printed tables were available. You should determine a value appropriate for your particular study.

## Report Options – Decimal Places

### Decimals in Means

Specify the number of decimal places to use when displaying means and effects.

### Decimals in Mean Squares

Specify the number of decimal places to use when displaying mean squares.

# Residual Plot, Effects Plot, and Means Plots Tabs

These options control the appearance of the two probability plots and the means plots that may be generated. More details on their interpretation will be contained in annotated output section presented later.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Y Scaling

Indicate whether the vertical scaling on all means plots should uniform across all plots.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## Plot Settings

### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Probability Plot procedure.

### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

### Connect Line(s)

Indicate whether to connect the means from the same factor with a line.

### Plot Settings - Legend

#### Show Legend

Specify whether to show the legend.

#### Legend Text

Enter text here for the legend title.

### Titles

#### Plot Title

This is the text of the title. Press the button on the right of the field to specify the font of the text.

## Symbols Tab

This section specifies the plot symbols.

### Plotting Symbols

#### Group 1-15

Specifies the plotting symbols used for each of the first fifteen groups.

## Storage Tab

The residuals calculated for each row may be stored on the current database for further analysis. This option lets you designate where to store the residuals.

### Data Storage Variables

#### Residuals

If a variable is specified here, the residuals are automatically stored in that variable. Note that any previous values in the variable are automatically replaced.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running the Analysis of a Two-Level Design

This section presents an example of how to analyze data using this program. We will analyze the three-factor experiment given on page 320 of Box and Hunter (1978). These data are the results of a pilot plant study conducted to investigate the influence of temperature, concentration, and catalyst on chemical yield. The data are contained in the BOX320 database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Analysis of Two-Level Designs window.

**1 Open the Box320 dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Box320.s0**.
- Click **Open**.

**2 Open the Analysis of Two-Level Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Analysis of Two-Level Designs**. The Analysis of Two-Level Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Analysis of Two-Level Designs window, select the **Variables tab**.
- Double-click in the **Response Variable** box. This will bring up the variable selection window.
- Select **Yield** from the list of variables and then click **Ok**. "Yield" will appear in the Response Variables box.
- Double-click in the **Factor Variables** box. This will bring up the variable selection window.
- Select **Temp, Concentration, Catalyst** from the list of variables and then click **Ok**. "Temp-Catalyst" will appear in the Factor Variables box.

**4 Specify the reports.**

- On the Analysis of Two-Level Designs window, select the **Reports tab**.
- Select **All** in the Confounding Size list box.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Design Information Section

**Design Information Section**

**Input Data**

| | |
|---|---|
| Response: | Yield |
| Rows: | 16 |
| Reps: | 2 |
| Blocks: | None |

| Factor Symbol | Factor Name | Level One | Level Two |
|---|---|---|---|
| A(1) | Temp | 160 | 180 |
| B(2) | Concentration | 20 | 40 |
| C(3) | Catalyst | 1 | 2 |

**Design**

2/1 replication of 3 factors.

This section describes the experimental design of the data.

## Confounding / Alias Section

**Confounding / Alias Section**

| Term No. | Terms Confounded |
|---|---|
| 1 | A |
| 2 | B |
| 3 | AB |
| 4 | C |
| 5 | AC |
| 6 | BC |
| 7 | ABC |

This section reports confounding and aliasing information for each term (degree of freedom). In the present example, a complete replication is given so there is no confounding. Hence, each degree of freedom is associated with only one term in the analysis of variance model.

# Means and Effects Section

**Means and Effects Section**

| Term No. | Term Symbol | Mean - | Mean + | Estimated Effect | Standard Error |
|---|---|---|---|---|---|
| 0 | Grand Mean | | | 64.25 | 0.71 |
| 1 | A (Temp) | 52.75 | 75.75 | 23.00 | 1.41 |
| 2 | B (Concentration) | 66.75 | 61.75 | -5.00 | 1.41 |
| 3 | AB | 63.50 | 65.00 | 1.50 | 1.41 |
| 4 | C (Catalyst) | 63.50 | 65.00 | 1.50 | 1.41 |
| 5 | AC | 59.25 | 69.25 | 10.00 | 1.41 |
| 6 | BC | 64.25 | 64.25 | 0.00 | 1.41 |
| 7 | ABC | 64.00 | 64.50 | 0.50 | 1.41 |

This section reports on the estimation of the effects for each degree of freedom.

### Term No.

This is an arbitrary identification number assigned to each degree of freedom in the model. This number is needed to correctly specify terms to be pooled as MSE.

### Term Symbol

This is the letter that is assigned to each factor. Since we have three factors in this database, three letters (A, B, and C) are used. The names of the variables associated with a given letter are shown in parentheses.

### Mean -

The average of all observations having the low value (-1) for this term. If you think of the low value as -1 and the high value as +1, then interaction terms (formed by multiplication) will also have only two possible values, -1 and +1.

### Mean +

The average of all observations having the high value (+1) for this term. If you think of the low value as -1 and the high value as +1, then interaction terms (formed by multiplication) will also have only two possible values, -1 and +1.

### Estimated Effect

The estimated effect value. This is equal to (Mean +) - (Mean -).

### Standard Error

The estimated standard error of the above effect value. Note that this standard error only depends on the MSE, so it is constant for all terms. Remember, this value is not calculated from individual groups of data but from the MSE!

## Sorted Means and Effects Section

**Sorted Means and Effects Section**

| Term No. | Term Symbol | Mean - | Mean + | Estimated Effect | Standard Error |
|---|---|---|---|---|---|
| 0 | Grand Mean | | | 64.25 | 0.71 |
| 1 | A (Temp) | 52.75 | 75.75 | 23.00 | 1.41 |
| 5 | AC | 59.25 | 69.25 | 10.00 | 1.41 |
| 2 | B (Concentration) | 66.75 | 61.75 | -5.00 | 1.41 |
| 4 | C (Catalyst) | 63.50 | 65.00 | 1.50 | 1.41 |
| 3 | AB | 63.50 | 65.00 | 1.50 | 1.41 |
| 7 | ABC | 64.00 | 64.50 | 0.50 | 1.41 |
| 6 | BC | 64.25 | 64.25 | 0.00 | 1.41 |

This is a sorted version of the report presented in the last section. The report is sorted by the absolute value of the Estimated Effect. This report is used with the Probability Plot of Effects to pick those effects that are not large enough to be important and thus can be pooled (averaged) into the MSE.

## Analysis of Variance Section

**Analysis of Variance Section**

| Term No. | Term Symbol | DF | Mean Square | F-Ratio | Prob Level | Statistically Significant |
|---|---|---|---|---|---|---|
| 1 | A (Temp) | 1 | 2116.0000 | 264.50 | 0.000000 | Yes |
| 2 | B (Concentration) | 1 | 100.0000 | 12.50 | 0.007670 | Yes |
| 3 | AB | 1 | 9.0000 | 1.13 | 0.319813 | No |
| 4 | C (Catalyst) | 1 | 9.0000 | 1.13 | 0.319813 | No |
| 5 | AC | 1 | 400.0000 | 50.00 | 0.000105 | Yes |
| 6 | BC | 1 | 0.0000 | 0.00 | 1.000000 | No |
| 7 | ABC | 1 | 1.0000 | 0.13 | 0.732810 | No |
| | Error | 8 | 8.0000 | | | |
| | Total | 15 | 2699.0000 | | | |

This section presents the analysis of variance table.

### Term No.

The identification number of this degree of freedom.

### Term Symbol

This is the symbol of this term. Refer to the Confounding / Alias Section for a list of all main effects and interactions associated with this term.

### DF

The degrees of freedom. The number of observations used by the corresponding model term.

### Mean Square

An estimate of the variation in the response accounted for by this term. The sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and mean square error (MSE). This F-ratio tests the statistical significance of the effects associated with this term.

### Prob Level

The probability of obtaining an F-ratio larger than that obtained by the analysis.

### Statistically Significant

If the probability level is less than the value of alpha that was set, the term is designated as being statistically significant (Yes). If it is not less than alpha, the term is not statistically significant (No).

## Two-Way Tables of Means and Effects

**Means and Effects of Concentration by Temp**

| Concentration | Temp 160 | Temp 180 | Effect | Overall |
|---|---|---|---|---|
| 20 | 56.00 | 77.50 | 21.50 | 66.75 |
| 40 | 49.50 | 74.00 | 24.50 | 61.75 |
| Effect | -6.50 | -3.50 | 1.50 | -5.00 |
| Overall | 52.75 | 75.75 | 23.00 | 64.25 |

This report presents the two-way interaction means and effects. One report is displayed for each of the possible two-way interactions.

The four means in the upper left-hand corner (56.0, 77.5, 49.5, 74.0) are the individual means. For example, 56.0 is the average of all rows in which the Temp was 160 and Concentration was 20.

The fourth column (Effect) is the estimated effect for that row. In the first, second, and fourth rows, this is the difference between the two previous columns. For example, 77.50-56.00 = 21.50. In the third row (Effect), the value (the effect of the interaction) is calculated as the difference between the two previous columns divided by 2. In this example, ((-3.50) - ( -6.50))/2 = 1.50.

The fifth column (Overall) gives the mean or effect for this row averaged across all column values. Thus, 66.75 is the average of all rows in which the value of Concentration was 20. Finally, 64.25 is the average of all rows in the experiment.

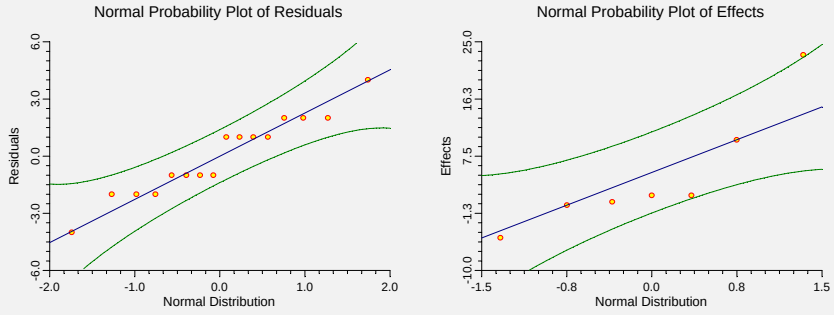
## Three-Way Tables of Means

**Means and Effects of Concentration by Temp by Catalyst**

| Concentration | Temp 160 Catalyst 1 | Temp 160 Catalyst 2 | Temp 180 Catalyst 1 | Temp 180 Catalyst 2 |
|---|---|---|---|---|
| 20 | 60.00 | 52.00 | 72.00 | 83.00 |
| 40 | 54.00 | 45.00 | 68.00 | 80.00 |

This report presents a three-way table of the means. For example, 60.00 is the average of all rows in which Temp was 160, Concentration was 20, and Catalyst was 1.

## Probability Plots Section

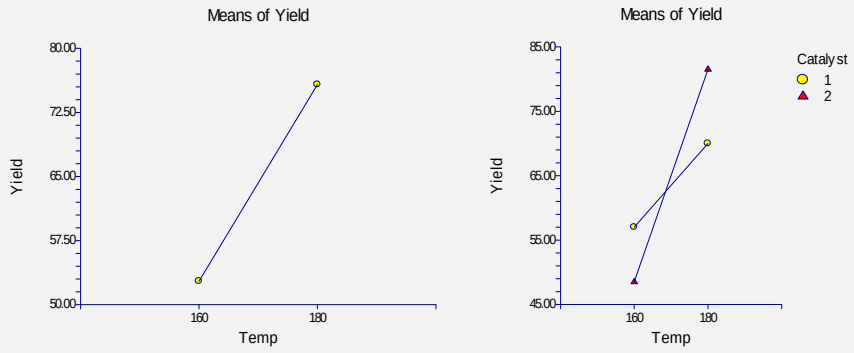



A normal probability plot of the residuals is supplied to allow you to study the distribution of the residuals. This plot will not be displayed if the residuals are all zero (which often occurs in designs like fractional factorials).

A normal probability plot of the effects is supplied to allow you to consider the relative sizes of the effects. If all terms are non-significant (and hence come from the normal distribution), these effects should fall along a straight line. When some of the effects are significant, they will fall off this line. The plot is useful for visually interpreting designs that do not supply an explicit estimate of the experimental error variance (such as fractional factorial designs).

In our example, the first six points seem to fall along a straight line, while the final point falls off this line. This term is associated with Temp (factor A), as you can see from the Means and Effects report.

## Means Plots Section




These plots display the means for all one-way and two-way interaction terms.

# Example 2 – Analysis of a Two-Level Design

This section presents another example of how to analyze data using this program. We will analyze an eight-factor experiment given on page 402 of Box and Hunter (1978). These data are the results of an injection molding study. The data are contained in the BOX402 database.

To run this example, open the BOX402 database and load the completed template **Example2** from the Template tab of the Analysis of Two-Level Designs window. Running this template will yield the following results.

## Design Information Section

**Design Information Section**

**Input Data**

| | |
|---|---|
| Response: | Shrinkage |
| Rows: | 16 |
| Reps: | 1 |
| Blocks: | None |

| Factor Symbol | Factor Name | Level One | Level Two |
|---|---|---|---|
| A(1) | MoldTemp | 1 | 2 |
| B(2) | Moisture | 1 | 2 |
| C(3) | HoldPressure | 1 | 2 |
| D(4) | Thickness | 1 | 2 |
| E(5) | BoosterPressure | 1 | 2 |
| F(6) | CycleTime | 1 | 2 |
| G(7) | GateSize | 1 | 2 |
| H(8) | ScrewSpeed | 1 | 2 |

**Design**
1/16 replication of 8 factors.

**Defining Contrast**
i = BCDE = ACDF = ABEF = ABCG = ADEG = BDFG = CEFG = ABDH = ACEH = BCFH = DEFH = CDGH = BEGH = AFGH = ABCDEFGH

**Design Construction**
Full model of the factors [ A B C D ].
The remaining factors are aliased with interactions
of this reduced model as follows:
E=BCD F=ACD G=ABC H=ABD

Notice the Defining Contrast and the Design Construction reports.

## Confounding / Alias Section

**Confounding / Alias Section**

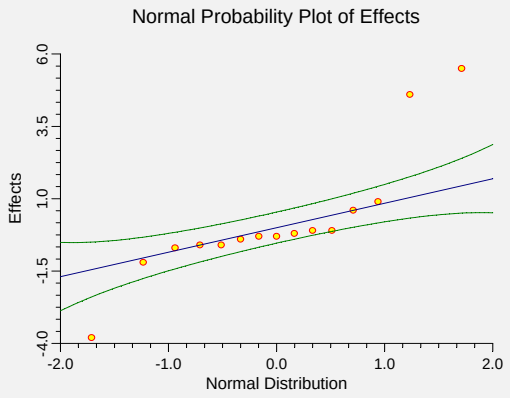
| Term No. | Terms Confounded |
|---|---|
| 1 | A+CDF+BEF+BCG+DEG+BDH+CEH+FGH |
| 2 | B+CDE+AEF+ACG+DFG+ADH+CFH+EGH |
| 3 | AB+EF+CG+DH |
| 4 | C+BDE+ADF+ABG+EFG+AEH+BFH+DGH |
| 5 | AC+DF+BG+EH |
| 6 | BC+DE+AG+FH |
| 7 | G+ABC+ADE+BDF+CEF+CDH+BEH+AFH |
| 8 | D+BCE+ACF+AEG+BFG+ABH+EFH+CGH |
| 9 | AD+CF+EG+BH |
| 10 | BD+CE+FG+AH |
| 11 | H+ABD+ACE+BCF+DEF+CDG+BEG+AFG |
| 12 | CD+BE+AF+GH |
| 13 | F+ACD+ABE+BDG+CEG+BCH+DEH+AGH |
| 14 | E+BCD+ABF+ADG+CFG+ACH+DFH+BGH |
| 15 | AE+BF+DG+CH |

This section reports confounding and aliasing information for each term (degree of freedom). Note that in this design, no two-way interactions are confounded with any of the main effects. Note, however, the all two-way interactions are confounded with each other.

## Sorted Means and Effects and Probability Plot Sections

**Means and Effects Section**

| Term No. | Term Symbol | Mean - | Mean + | Estimated Effect | Standard Error |
|---|---|---|---|---|---|
| 0 | Grand Mean | | | 19.75 | 0.30 |
| 4 | C (HoldPressure) | 17.00 | 22.50 | 5.50 | 0.59 |
| 15 | AE | 17.45 | 22.05 | 4.60 | 0.59 |
| 14 | E (BoosterPressure) | 21.65 | 17.85 | -3.80 | 0.59 |
| 11 | H (ScrewSpeed)* | 20.35 | 19.15 | -1.20 | 0.59 |
| 5 | AC* | 19.30 | 20.20 | 0.90 | 0.59 |
| 1 | A (MoldTemp)* | 20.10 | 19.40 | -0.70 | 0.59 |
| 7 | G (GateSize)* | 19.45 | 20.05 | 0.60 | 0.59 |
| 3 | AB* | 20.05 | 19.45 | -0.60 | 0.59 |
| 10 | BD* | 20.05 | 19.45 | -0.60 | 0.59 |
| 9 | AD* | 19.95 | 19.55 | -0.40 | 0.59 |
| 8 | D (Thickness)* | 19.90 | 19.60 | -0.30 | 0.59 |
| 12 | CD* | 19.90 | 19.60 | -0.30 | 0.59 |
| 6 | BC* | 19.85 | 19.65 | -0.20 | 0.59 |
| 2 | B (Moisture)* | 19.80 | 19.70 | -0.10 | 0.59 |
| 13 | F (CycleTime)* | 19.80 | 19.70 | -0.10 | 0.59 |




From the probability plot you can see that three of the effects fall outside the range that would be expected if all effects come from the normal distribution. By looking at the Sorted Means and Effects report, we see that these three terms are numbers 4, 15, and 14. Hence, we decided to pool the rest of the terms to form an estimate of the experimental error variance (MSE) and rerun the program. We add the text 1,2,3,5,6,7,8,9,10,11,12,13 to the Pooled Terms option and rerun. The following analysis of variance table is produced. Note that without pooling these terms, the error DF would have been zero and no F-Ratios would have been generated.

## Analysis of Variance Section

**Analysis of Variance Section**

| Term No. | Term Symbol | DF | Mean Square | F-Ratio | Prob Level | Statistically Significant |
|---|---|---|---|---|---|---|
| 1 | A (MoldTemp)* | 1 | 1.9600 | 1.39 | 0.260707 | No |
| 2 | B (Moisture)* | 1 | 0.0400 | 0.03 | 0.868897 | No |
| 3 | AB* | 1 | 1.4400 | 1.02 | 0.331610 | No |
| 4 | C (HoldPressure) | 1 | 121.0000 | 86.02 | 0.000001 | Yes |
| 5 | AC* | 1 | 3.2400 | 2.30 | 0.154993 | No |
| 6 | BC* | 1 | 0.1600 | 0.11 | 0.741747 | No |
| 7 | G (GateSize)* | 1 | 1.4400 | 1.02 | 0.331610 | No |
| 8 | D (Thickness)* | 1 | 0.3600 | 0.26 | 0.622099 | No |
| 9 | AD* | 1 | 0.6400 | 0.45 | 0.512767 | No |
| 10 | BD* | 1 | 1.4400 | 1.02 | 0.331610 | No |
| 11 | H (ScrewSpeed)* | 1 | 5.7600 | 4.09 | 0.065871 | No |
| 12 | CD* | 1 | 0.3600 | 0.26 | 0.622099 | No |
| 13 | F (CycleTime)* | 1 | 0.0400 | 0.03 | 0.868897 | No |
| 14 | E (BoosterPressure) | 1 | 57.7600 | 41.06 | 0.000034 | Yes |
| 15 | AE | 1 | 84.6400 | 60.17 | 0.000005 | Yes |
| | Error | 12 | 1.4067 | | | |
| | Total | 15 | 280.2800 | | | |

Terms marked with an '*' have been pooled to form the error term.

Now we see that HoldPressure, BoosterPressure, and at least one of the two-way interactions AE+BF+DG+CH are significant.

## Chapter 214

# Repeated Measures Analysis of Variance

## Introduction

This procedure performs an analysis of variance on repeated measures (within-subject) designs using the general linear models approach. The experimental design may include up to three between-subject terms as well as three within-subject terms. Box's M and Mauchly's tests of the assumptions about the within-subject covariance matrices are provided. Geisser-Greenhouse, Box, and Huynh-Feldt corrected probability levels on the within-subject F tests are given along with the associated test power.

Repeated measures designs are popular because they allow a subject to serve as their own control. This improves the precision of the experiment by reducing the size of the error variance on many of the F-tests, but additional assumptions concerning the structure of the error variance must be made.

This procedure uses the general linear model (GLM) framework to perform its calculations. Identical results can be achieved by using the GLM ANOVA program. The user input of this procedure is simply the GLM panel modified to allow a more direct specification of a repeated-measures model. We refer you to the GLM ANOVA chapter for details on the calculations and interpretations of analysis of variance. We will concentrate here on providing information specific to repeated measures analysis.

## An Example

This section will give an example of a repeated-measures experiment. An experiment was conducted to study the effects of four drugs upon reaction time to a set of tasks using five subjects.

| Subject | Drug 1 | Drug 2 | Drug 3 | Drug 4 |
|---|---|---|---|---|
| 1 | 30 | 28 | 16 | 34 |
| 2 | 14 | 18 | 10 | 22 |
| 3 | 24 | 20 | 18 | 30 |
| 4 | 38 | 34 | 20 | 44 |
| 5 | 26 | 28 | 14 | 30 |

# Discussion

One way of categorizing experimental designs is as *between subject* or *within subject.* Examples of between-subject designs are the common factorial designs in which the experimental units (the subjects) are assigned to separate treatment conditions. Usually, this assignment is done at random. The experimenter wants to know if the variability from subject to subject is smaller than the variability from treatment to treatment. The basic assumption is that the subjects are independent from one another.

Within-subject designs are those in which multiple measurements are made on the same individual. Because the response to stimuli usually varies less within an individual than between individuals, the within-subject variability is usually less than (or at most equal to) the between-subject variability. Reducing the underlying variability reduces the sample size which reduces cost.

## Disadvantages of Within-Subjects Designs

The main advantage of within-subjects designs is in the reduced variability that is achieved by controlling from differences from one subject to the next. There are several disadvantages to this type of design:

1. Practice effect. In some experiments, subjects systematically improve as they practice the task being studies. In other cases, subjects may systematically get worse as the get fatigued or bored with the experimental task. Note that only the treatment administered first is immune to practice effects. Hence, experimenters often make some effort to balance the number of subjects receiving each treatment first.
2. Carryover effect. In many drug studies, it is important to “wash out” one drug completely before the next drug is administered. Otherwise, the influence of the first drug carries over into the response to the second drug. Note that practice effects refer to a general change in response because the task is repeated, but carryover effects refer to specific, lasting effects of a particular treatment.
3. Statistical analysis. The statistical model that justifies the analysis is very restrictive since the individual responses must have certain mathematical properties. Also, missing responses are much more difficult to deal with in this case.
4. Generalizability. Experimenters assume that differences between treatments are design independent. That is, if a completely random design was constructed, the same treatment differences would be observed. This is not always the case.

Even in the face of all these disadvantages, repeated measures (within-subject) designs are popular in many areas of research. It is important that you recognize these problems going in, rather than learning of them later after the experiment has been conducted.

# Assumptions

The following assumptions are made when using the *F* test to analyze a factorial experimental design.

1. The response variable is continuous.
2. The residuals follow the normal probability distribution with mean equal to zero and constant variance.

3. The subjects are independent. Since in a within-subject design, responses coming from the same subject are not usually independent, assumption three must be modified for responses within a subject. The independence between subjects is still assumed.
4. The within-subject covariance matrices are equal for all between-subject groups. In this type of experiment, the repeated measurements on a subject may be thought of as a multivariate response vector having a certain covariance structure. This assumption states that these covariance matrices are constant from group to group. This assumption is tested by Box's M test. Of course, this assumption unnecessary in the single-group design.
5. All of the within-subject covariance matrices are circular. One way of defining circularity is that the variances of differences between any two measurements within a subject are constant. Since responses that are close together in time often have a higher correlation than those that are far apart, it is common for this assumption to be violated. This assumption is tested by Mauchly's test and be studying the values of epsilon (defined below). The circularity assumption is not necessary when only two repeated measures are made.

The program provides formal tests of these assumptions. However, these tests have their own assumptions which also may be violated, so a more common strategy is to assume that the circularity is violated and take appropriate action. ***NCSS*** does this for you automatically.

# Technical Details

Other than reformatting the input panel, the main difference between this procedure and the GLM procedure is the inclusion of the Geisser-Greenhouse correction and associated tests of assumptions. Because of this, we will present only those results here. You can obtain a more general overview of analysis of variance in the One-Way Analysis of Variance and General Linear Models chapters.

## Covariance Matrix Assumptions

The covariance matrix for a design with *m* subjects and *k* measurements per subject may be represented as

$$\Sigma = \left[\sigma_{ij}\right]$$

Valid F tests in a repeated-measures design require that the covariance matrix is a type *H* matrix. A type *H* matrix has the *circularity* property that

$$\Sigma = A + A' + \lambda I_k$$

where $I_k$ is the identity matrix of order *k* and $\lambda$ is a constant.

This property may also be defined as

$$\sigma_{ii} + \sigma_{jj} - 2\sigma_{ij} = 2\lambda$$

One type of matrix that has this property is one which has *compound symmetry*. A matrix with this property has all elements on the main diagonal equal and all elements off the main diagonal equal. An example of a covariance matrix with compound symmetry is

$$\begin{bmatrix} 9 & 2 & 2 & 2 \\ 2 & 9 & 2 & 2 \\ 2 & 2 & 9 & 2 \\ 2 & 2 & 2 & 9 \end{bmatrix}$$

An example of a type *H* matrix which does not have compound symmetry is

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 4 & 5 \\ 3 & 6 & 5 & 6 \\ 4 & 5 & 8 & 7 \\ 5 & 6 & 7 & 10 \end{bmatrix}$$

Note that if the diagonal elements are equal, which implies that the variation within each subject is constant, a type *H* matrix must have compound symmetry.

## Epsilon

Epsilon is a measure of the extent to which a covariance matrix departs from circularity. It was developed by Box (see Winer(1991) or Kirk (1982)) and is estimated by

$$\hat{\varepsilon} = \frac{k^2\left(\sum_{i=1}^{k}\frac{s_{ii}}{k} - \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{s_{ij}}{k^2}\right)^2}{(k-1)\left[\sum_{i=1}^{k}\sum_{j=1}^{k}s_{ij}^2 - 2k\sum_{j=1}^{k}\left(\sum_{i=1}^{k}\frac{s_{ji}}{k}\right)^2 + k^2\left(\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{s_{ij}}{k^2}\right)^2\right]}$$

where the estimated covariance matrix is given by

$$\hat{\Sigma} = \left[s_{ij}\right]$$

and *k* is the number of levels of the within subject factor.

For two- and three-way interaction terms, epsilon is estimated by

$$\hat{\varepsilon} = \frac{\left(\sum_{i=1}^{r} z_{ii}\right)^2}{r\sum_{i=1}^{r}\sum_{j=1}^{r} z_{ij}^2}$$

where $Z = CSC'$ and *C* is a contrast matrix appropriate for testing that interaction.

This estimate of epsilon is biased, especially for large values of epsilon. To correct for this bias, Huynh and Feldt developed another estimate of epsilon, which is calculated as follows

$$\tilde{\varepsilon} = Min\left[\frac{N(k-1)\hat{\varepsilon} - 2}{(k-1)\left[N - g - (k-1)\hat{\varepsilon}\right]}, 1\right]$$

where *N* is the total number of subjects and *g* is the number of levels of the between factors.

The range of epsilon is

$$\frac{1}{k-1} \le \varepsilon \le 1$$

When $\varepsilon = 1$, the matrix is circular. When $\varepsilon = \frac{1}{k-1}$, the matrix differs maximally from circularity.

Box's estimator tends to underestimate epsilon and the Huynh-Feldt estimate tends to overestimate it. Simulation studies have found Box's estimate to be the one that should be used to adjust the F tests.

## Geisser-Greenhouse Adjustment

All F ratios of within subject factors and interactions require the assumption that the covariance matrix is of type *H* in order for the F ratio to follow the F distribution with degrees of freedom *df1* and *df2*. When the covariance matrix is not of type *H*, Geisser and Greenhouse suggested that the distribution of the F ratio be approximated by an F distribution with degrees of freedom $\varepsilon(df1)$ and $\varepsilon(df2)$ where $\varepsilon$ is set at its minimum, that is, $\varepsilon = \frac{1}{k-1}$.

Box suggested that rather than use the minimum to adjust the degrees of freedom, $\varepsilon$ should be set at the Geisser-Greenhouse value, $\hat{\varepsilon}$. Realizing the $\hat{\varepsilon}$ is biased, Huynh and Feldt suggested that $\tilde{\varepsilon}$ be used. Simulation studies have shown that using Box's adjustment consistently gives the most accurate significance levels.

## Mauchly's Test of Compound Symmetry

Mauchly (1940) developed a test to determine if a sample covariance matrix has compound symmetry. The formulas for Mauchly's test statistic *W*, given in Huynh and Feldt (1970), are as follows

$$W = |CSC'| / \left(\text{trace}CSC'/p\right)^p$$

$$\chi^2_{p(p+1)/2-1} = -(N-g)\left(1 - \frac{2p^2+p+2}{6p(N-g)}\right)\ln(W)$$

where *g* is the number of groups, *N* is the number of subjects, *C* is a contrast matrix with *p* rows suitable for testing a main effect or interaction, *S* is a *k*-by-*k* matrix of the pooled group covariances. Note that usually, *p* equals the degrees of freedom of the corresponding term.

# Data Structure

The data must be entered in a format that puts the response values in one variable and the values of each of the factors in other variables. We will first present an example of a single-group repeated measures design followed by an example of a design with one between factor and one within factor.

## Single-Group Repeated Measures Design Example – REACTION Database

The experiment described in this example was conducted to study the effects of four drugs upon reaction time to a set of tasks. The five subjects were given extensive training in the tasks prior to the experiment so that there would be no carryover (learning) effect from one trial to the next. The five subjects were chosen at random from the population of interest.

The order in which the drugs were administered was randomized separately for each subject. A sufficient time was allowed between trials to wash out the effect of the previous drug before administering the next drug. The results of this experiment are recorded in the REACTION database. This design is often referred to as a *randomized block design*.

**REACTION dataset**

| Row | Person | Drug | Test |
|---|---|---|---|
| 1 | 1 | 1 | 30 |
| 2 | 1 | 2 | 28 |
| 3 | 1 | 3 | 16 |
| 4 | 1 | 4 | 34 |
| 5 | 2 | 1 | 14 |
| 6 | 2 | 2 | 18 |
| 7 | 2 | 3 | 10 |
| 8 | 2 | 4 | 22 |
| 9 | 3 | 1 | 24 |
| 10 | 3 | 2 | 20 |

| Row | Person | Drug | Test |
|---|---|---|---|
| 11 | 3 | 3 | 18 |
| 12 | 3 | 4 | 30 |
| 13 | 4 | 1 | 38 |
| 14 | 4 | 2 | 34 |
| 15 | 4 | 3 | 20 |
| 16 | 4 | 4 | 44 |
| 17 | 5 | 1 | 26 |
| 18 | 5 | 2 | 28 |
| 19 | 5 | 3 | 14 |
| 20 | 5 | 4 | 30 |

## Heart Rate Data - EXERCISE Database

The following dataset is an example of a one between-factor and one within-factor repeated measures design. An experiment was conducted to study the effects of exercise on heart rate. The subjects were randomly divided into three groups of six subjects each. The first group did not have a regular exercise plan. The second group exercised once a week. The third group exercised daily. Each subject's heart rate was recorded when the experiment began, at the end of ten weeks, and at the end of twenty weeks. These data are stored in a database called EXERCISE.

**EXERCISE dataset**

| Row | Exercise | Subj | Time | Heart Rate |
|---|---|---|---|---|
| 1 | 0 - None | 1 | 0 | 87 |
| 2 | 0 - None | 1 | 10 | 77 |
| 3 | 0 - None | 1 | 20 | 84 |
| 4 | 0 - None | 2 | 0 | 67 |
| 5 | 0 - None | 2 | 10 | 65 |
| 6 | 0 - None | 2 | 20 | 62 |
| 7 | 0 - None | 3 | 0 | 55 |
| 8 | 0 - None | 3 | 10 | 52 |
| 9 | 0 - None | 3 | 20 | 58 |
| 10 | 0 - None | 4 | 0 | 66 |
| 11 | 0 - None | 4 | 10 | 70 |
| 12 | 0 - None | 4 | 20 | 65 |
| 13 | 0 - None | 5 | 0 | 88 |
| 14 | 0 - None | 5 | 10 | 82 |
| 15 | 0 - None | 5 | 20 | 85 |
| 16 | 0 - None | 6 | 0 | 75 |
| 17 | 0 - None | 6 | 10 | 72 |
| 18 | 0 - None | 6 | 20 | 79 |
| 19 | 1 - Weekly | 7 | 0 | 84 |
| 20 | 1 - Weekly | 7 | 10 | 78 |
| 21 | 1 - Weekly | 7 | 20 | 74 |
| 22 | 1 - Weekly | 8 | 0 | 78 |
| 23 | 1 - Weekly | 8 | 10 | 72 |
| 24 | 1 - Weekly | 8 | 20 | 68 |
| 25 | 1 - Weekly | 9 | 0 | 64 |
| 26 | 1 - Weekly | 9 | 10 | 53 |
| 27 | 1 - Weekly | 9 | 20 | 54 |

| Row | Exercise | Subj | Time | Heart Rate |
|---|---|---|---|---|
| 28 | 1 - Weekly | 10 | 0 | 73 |
| 29 | 1 - Weekly | 10 | 10 | 68 |
| 30 | 1 - Weekly | 10 | 20 | 63 |
| 31 | 1 - Weekly | 11 | 0 | 84 |
| 32 | 1 - Weekly | 11 | 10 | 77 |
| 33 | 1 - Weekly | 11 | 20 | 74 |
| 34 | 1 - Weekly | 12 | 0 | 55 |
| 35 | 1 - Weekly | 12 | 10 | 53 |
| 36 | 1 - Weekly | 12 | 20 | 52 |
| 37 | 2 - Daily | 13 | 0 | 72 |
| 38 | 2 - Daily | 13 | 10 | 55 |
| 39 | 2 - Daily | 13 | 20 | 53 |
| 40 | 2 - Daily | 14 | 0 | 83 |
| 41 | 2 - Daily | 14 | 10 | 72 |
| 42 | 2 - Daily | 14 | 20 | 69 |
| 43 | 2 - Daily | 15 | 0 | 75 |
| 44 | 2 - Daily | 15 | 10 | 63 |
| 45 | 2 - Daily | 15 | 20 | 65 |
| 46 | 2 - Daily | 16 | 0 | 55 |
| 47 | 2 - Daily | 16 | 10 | 49 |
| 48 | 2 - Daily | 16 | 20 | 51 |
| 49 | 2 - Daily | 17 | 0 | 83 |
| 50 | 2 - Daily | 17 | 10 | 76 |
| 51 | 2 - Daily | 17 | 20 | 72 |
| 52 | 2 - Daily | 18 | 0 | 63 |
| 53 | 2 - Daily | 18 | 10 | 54 |
| 54 | 2 - Daily | 18 | 20 | 55 |

## Missing Values

There are two kinds of unbalance that can occur in repeated-measures designs. First, in multi-group designs, there may be a different number of subjects in each group. This type of unbalance causes no problems in the F-tests. Second, some individuals may not have had all measurements. When this occurs, the program makes the additional assumption that the within-subject sample effects sum to zero. Every effort should be made to avoid missing values because of the additional assumptions that must be made. However, even when data are missing, meaningful conclusions can be drawn.

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

This panel specifies the variables used in the analysis.

### Response Variables

#### Response Variable(s)

Specifies the response (measurement) variable to be analyzed. Only one variable is needed for an analysis. If you specify more than one variable, a separate analysis is run for each variable.

Note that only one measurement is entered on each row. Hence, a repeated measurement design with five measurements per subject will require five rows per subject on the database.

### Subject Variable

#### Subject Variable

A single subject factor is required. In a repeated measures design, the subjects are categorized into one or more mutually exclusive groups and each subject is measured two or more times. This variable identifies the subject associated with the measurement. Each subject must be identified with a unique name or number.

### Between Factors

#### Between Factor (1-3)

From zero to three between factor variables may be specified. A Between Factor specifies a way of categorizing the subjects. Examples of between factors are gender, age groups, and blood type. If none are specified, a single-group repeated-measures analysis is run.

Values in the variable may be either numeric or text.

#### Random

This option specifies whether the factor is fixed or random. These options control the denominator terms of the F-ratio values.

A **fixed** factor includes all possible levels, like male and female for gender, includes representative values across the possible range of values, like low, medium, and high blood pressure, or includes a set of values to which inferences will be limited, like New York, California, and Maryland.

A **random** factor is one in which the chosen levels represent a random sample from the population of values. For example, you might select four classes from the hundreds in your state or you might select ten batches from an industrial process. The key is that a random sample is chosen. In ***NCSS***, a random factor is “crossed” with other random and fixed factors. Two factors are crossed when each level of one includes all levels of the other.

## Within Factors

### Within Factor (1-3)

At least one within factor variable must be specified. A Within Factor specifies a way of categorizing the measurements made on each subject. For example, a measurement may be made at one week, two weeks, and three weeks. Weeks would be the within factor.

### Random

This option specifies whether the factor is fixed or random. Usually, within factors are fixed.

## Model Specification

This section specifies the experimental design model.

### Which Model Terms

A design in which all main effect and interaction terms are included is called a saturated model. Occasionally, it is useful to omit various interaction terms from the model—usually because some data values are missing. This option lets you specify which interactions to keep.

The options included here are:

- **Full Model. Use all terms.**

  The complete, saturated model is analyzed. All reports will be generated when this option is selected.

- **Full model except subject interactions combined with error.**

  Some authors recommend pooling the interactions involving the subject factor into one error term to achieve more error degrees of freedom and thus more power in the F-tests. This option lets you do this. Note that the Geisser-Greenhouse corrections are not made in this case.

- **Use the Custom Model given below.**

  This option indicates that you want the Custom Model (given in the next box) to be used.

- **Custom Model**

  When a custom model (see Which Model Terms above) is selected, you will enter the actual model here. If all main effects and interactions are desired, you can enter the word “ALL” here. For complicated designs, it is usually easier to check the next option, Write Model in ‘Custom Model’ Field, and run the procedure. The appropriate model will be generated and placed in this box. You can then delete the terms you do not want.

  The model is entered using letters separated by the plus sign. For example, a one-between factor and one-within factor repeated-measures design would look like this:

  A+B(A)+C+AC+BC(A).

- **Write model in ‘Custom Model’ field.**

  When this option is checked, no analysis is performed when the procedure is run. Instead, a copy of the full model is stored in the Custom Model box. You can then delete selected terms from the model without having to enter all the terms you want to keep.

# Comparisons Tab

These panels specify the planned comparisons for the between and within factors.

## Between and Within Factor Planned Comparisons

### Comparison (1-3)

This option lets you specify individual comparisons for each factor. Comparisons are only valid for fixed factors. A comparison is formulated in terms of the means as follows:

$$C_i = \sum_{j=1}^{J} w_{ij} m_j$$

In this equation, there are J levels in the factor, the means for each level of the factor are denoted $m_i$, and $w_{ij}$ represents a set of J weight values for the $i^{th}$ comparison. The comparison value, $C_i$, is tested using a t-test. Note that if the $w_{ij}$ sum to zero across j, the comparison is called a "contrast" of the means.

Comparisons are specified by listing the weights. For example, suppose a factor has three levels. Further suppose that the first level represents a control group, the second a treatment at one dose, and the third a treatment at a higher dose. Three comparisons come to mind: compare each of the treatment groups to the control group and compare the two treatment groups to each other. These three comparisons would be

| | |
|---|---|
| Control vs. Treatment 1 | -1,1,0 |
| Control vs. Treatment 2 | -1,0,1 |
| Treatment 1 vs. Treatment 2 | 0,-1,1 |

You might also be interested in comparing the control group with the average of both treatment groups. The weights for this comparison would be -2,1,1.

When a factor is quantitative, it might be of interest to divide the response pattern into linear, quadratic, cubic, or other components. If the sample sizes are equal and the factor levels are equally spaced, these so-called components of trend may be studied by the use of simple contrasts. For example, suppose a quantitative factor has three levels: 5, 10, and 15. Contrasts to test the linear and quadratic trend components would be

| | |
|---|---|
| Linear trend | -1,0,1 |
| Quadratic trend | 1,-2,1 |

If the sample sizes for the groups are unequal (the design is unbalanced), adjustments must be made for the differing sample sizes.

***NCSS*** will automatically generate some of the more common sets of contrasts, or it will let you specify up to three custom contrasts yourself. The following common sets are designated by this option.

- **None**

  No comparisons are generated.

- **Standard Set**

  This option generates a standard set of contrasts in which the mean of the first level is compared to the average of the rest, the mean of the second group is compared to the average of those remaining, and so on.

  The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

  | | |
  |---|---|
  | -3,1,1,1 | Compare the first-level mean with the average of the rest. |
  | 0,-2,1,1 | Compare the second-level mean with the average of the rest. |
  | 0,0,-1,1 | Compare the third-level mean with the fourth-level mean. |

- **Polynomial**

  This option generates a set of orthogonal contrasts that allow you to test various trend components from linear up to sixth order. These contrasts are appropriate even if the levels are unequally spaced or the group sample sizes are unequal. Of course, these contrasts are only appropriate for data that are at least ordinal. Usually, you would augment the analysis of this type of data with a multiple regression analysis.

  The following example displays the type of contrasts generated by this option. Suppose there are four equally spaced levels in the factor and each group has two observations. The contrasts generated by this option are (scaled to whole numbers):

  | | |
  |---|---|
  | -3,-1,1,3 | Linear component. |
  | 1,-1,-1,1 | Quadratic component. |
  | -1,3,-3,1 | Cubic component. |

- **Linear Trend**

  This option generates a set of orthogonal contrasts and retains only the linear component. This contrast is appropriate even if the levels are unequally spaced and the group sample sizes are unequal. See Orthogonal Polynomials above for more detail.

- **Linear-Quadratic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first two (the linear and quadratic) are reported.

- **Linear-Cubic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first three are reported.

- **Linear-Quartic Trend**

  This option generates the complete set of orthogonal polynomials, but only the results for the first four are reported.

- **Each with First**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the first level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

| | |
|---|---|
| -1,1,0,0 | Compare the first- and second-level means. |
| -1,0,1,0 | Compare the first- and third-level means. |
| -1,0,0,1 | Compare the first- and fourth-level means. |

- **Each with Last**

  This option generates a set of nonorthogonal contrasts appropriate for comparing each of the remaining levels with the last level. The following example displays the type of contrast generated by this option. Suppose there are four levels (groups) in the factor. The contrasts generated by this option are:

| | |
|---|---|
| -1,0,0,1 | Compare the first- and fourth-level means. |
| 0,-1,0,1 | Compare the second- and fourth-level means. |
| 0,0,-1,1 | Compare the third- and fourth-level means. |

- **Custom**

  This option indicates that the contrasts entered in the three boxes below it should be used. The specification of these three boxes is described next.

### Custom (1-3)

These three boxes let you write a user-specified comparison by specifying the weights of that comparison. Note that there are no numerical restrictions on these coefficients. They do not even have to sum to zero. However, this is recommended. If the coefficients do sum to zero, the comparison is called a contrast. The significance tests anticipate that only one or two of these comparisons are to be run. If you run several, you should make some type of Bonferroni adjustment to your alpha value.

When you put in your own contrasts, you must be careful that you specify the appropriate number of weights. For example, if the factor has four levels, four weights must be specified, separated by commas. Extra weights are ignored. If too few weights are specified, the missing weights are set to zero.

These comparison coefficients designate weighted averages of the level-means that are to be statistically tested. The null hypothesis is that the weighted average is zero. The alternative hypothesis is that the weighted average is nonzero. The weights (comparison coefficients) are specified here.

As an example, suppose you want to compare the average of the first two levels with the average of the last two levels in a six-level factor. You would enter “-1,-1,0,0,1,1.”

As a second example, suppose you want to compare the average of the first two levels with the average of the last three levels in a six-level factor. The contrast would be “-3,-3,0,2,2,2."

Note that in each case, we have used weights that sum to zero. This is why we could not use ones in the second example.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Reports

### EMS Report ... Means Report

Specify whether to display the indicated reports.

## Select Plots

### Means Plot(s) and Subject Plot

Specify whether to display the indicated plots.

## Report Options

### F-Test Alpha

The value of alpha for the statistical tests and power analysis. Usually, this number will range from 0.10 to 0.001. A common choice for alpha is 0.05, but this value is a legacy from the age before computers when only printed tables were available. You should determine a value appropriate for your particular study.

### Assumptions Alpha

This option specifies the value of alpha used in the tests of assumptions: Box's M test and Mauchly's test. Most statisticians recommend that these preliminary tests be carried out at a higher alpha (probability of rejecting a true null hypothesis) value such as 0.10 or 0.20.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, while the double precision will display thirteen-place accuracy.

### Variable Names

Indicate whether to display the variable names or the variable labels.

### Value Labels

Indicate whether to display the data values or their labels.

## Multiple Comparison Tests

### Bonferroni Test (All-Pairs) ... Tukey-Kramer Confidence Intervals

These options specify which MC tests and confidence intervals to display.

### Tests for Two-Factor Interactions

This option specifies whether multiple comparison tests are generated for two-factor interaction terms. When checked, the means of two-factor interactions will be tested by each active multiple comparison test. The multiple comparison test will treat the means as if they came from a single factor. For example, suppose factor A as two levels and factor B has three levels. The AB interaction would then have six levels. The active multiple comparison tests would be run on these six means.

Care must be used when interpreting multiple comparison tests on interaction means. Remember that the these means contain not only the effects of the interaction, but also the main effects of the two factors. Hence these means contain the combined effects of factor A, factor B, and the AB interaction. You cannot interpret the results as representing only the AB interaction.

## Multiple Comparison Tests – Options

### MC Alpha

Specifies the alpha value used by the multiple-comparison tests.

### MC Decimals

Specify how many decimals to display in the multiple comparison sections.

# Means Plot and Subject Plot Tabs

These options specify the plots of group means and subject's responses across time.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Y Scaling (Means Plot)

Specify the method for calculating the minimum and maximum along the vertical axis. *Separately* means that each plot is scaled independently. *Uniform* means that all plots use the overall minimum and maximum of the data. This option is ignored if a minimum or maximum is specified.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Lines

Click this box to connect the points for a particular factor. This makes it easier to spot patterns in the means.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, *{G}* is replaced by the appropriate factor name or subject variable name.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by appropriate names. *{G}* is replaced by the name of the subject variable. Press the button on the right of the field to specify the font of the text.

# Symbols Tab

These options specify the symbols used in the plots.

## Plotting Symbols

### Group (1-15)

The symbols used to represent the levels of a factor. Group 1 represents the first level, Group 2 represents the second level, and so on.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

## Example 1 – Running Repeated Measures ANOVA

This section presents an example of how to run an analysis of a typical repeated measures design with one between factor and one within factor. These data are contained in the EXERCISE database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Repeated Measures Analysis of Variance window.

**1 Open the EXERCISE dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **EXERCISE.S0**.
- Click **Open**.

**2 Open the Repeated Measures Analysis of Variance window.**

- On the menus, select **Analysis**, then **Analysis of Variance (ANOVA)**, then **Repeated Measures Analysis of Variance**. The Repeated Measures Analysis of Variance procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Repeated Measures Analysis of Variance window, select the **Variables tab**.
- Double-click in the **Response Variable(s)** box. This will bring up the variable selection window.
- Select **HeartRate** from the list of variables and then click **Ok**.
- Double-click in the **Subject Variable** box. This will bring up the variable selection window.
- Select **Subject** from the list of variables and then click **Ok**.
- Double-click in the **Between Factor 1** box. This will bring up the variable selection window.
- Select **Exercise** from the list of variables and then click **Ok**.
- Double-click in the **Within Factor 1** box. This will bring up the variable selection window.
- Select **Time** from the list of variables and then click **Ok**.

**4 Specify the planned comparison tests.**

- On the Repeated Measures Analysis of Variance window, select the **Comparisons tab**.
- Set the **Between Factor Planned Comparisons – Comparison 1** field to **Each with First**. This will generate the test of the no exercise group with the weekly exercise group and the no exercise group with the daily exercise group.

**5 Specify the multiple comparison tests.**

- On the Repeated Measures Analysis of Variance window, select the **Reports tab**.
- Check the **Tukey-Kramer Test** option of the Multiple Comparison Tests.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

We will now document this output, one section at a time.

## Expected Mean Squares Section

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A: Exercise | 2 | Yes | B(A) | S+csB+bcsA |
| B(A): Subject | 15 | No | S(ABC) | S+csB |
| C: Time | 2 | Yes | BC(A) | S+sBC+absC |
| AC | 4 | Yes | BC(A) | S+sBC+bsAC |
| BC(A) | 30 | No | S(ABC) | S+sBC |
| S(ABC) | 0 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

The expected mean square expressions are provided to show the appropriate error term for each factor. The correct error term for a factor is that term that is identical except for the factor being tested.

Note that in the repeated measures model, there are two error terms that are used: the between error labeled B(A) and the within error labeled BC(A).

### Source Term

The source of variation or term in the model.

### DF

The degrees of freedom, which is the number of observations used by this term.

### Term Fixed?

Indicates whether the term is fixed or random.

### Denominator Term

Indicates the term used as the denominator in the F-ratio. This is the error term for this term.

### Expected Mean Square

This expression represents the expected value of the corresponding mean square if the design was completely balanced. *S* represents the expected value of the mean square error (sigma). The uppercase letters represent either the adjusted sum of squared treatment means if the factor is fixed, or the variance component if the factor is random. The lowercase letter represents the number of levels for that factor, and *s* represents the number of replications of the experimental layout.

These EMS expressions are provided to determine the appropriate error term for each factor. The correct error term for a factor is that term whose EMS is identical except for the factor being tested.

# Analysis of Variance Table Section

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Exercise | 2 | 427.4445 | 213.7222 | 0.61 | 0.555040 | 0.133339 |
| B(A): Subject | 15 | 5234.556 | 348.9704 | | | |
| C: Time | 2 | 547.4445 | 273.7222 | 36.92 | 0.000000* | 1.000000 |
| AC | 4 | 191.4444 | 47.86111 | 6.45 | 0.000716* | 0.977632 |
| BC(A) | 30 | 222.4444 | 7.414815 | | | |
| S | 0 | | | | | |
| Total (Adjusted) | 53 | 6623.333 | | | | |
| Total | 54 | | | | | |

* Term significant at alpha = 0.05

### Source Term

The source of variation, which is the term in the model.

### DF

The degrees of freedom, which is the number of observations used by the corresponding model term.

### Sum of Squares

This is the sum of squares for this term. It is usually included in the ANOVA table for completeness, not for direct interpretation.

The sums of squares are calculated as follows. First, the sum of squares of each term that does not involve the subject factor is computed using the difference between two reduced models. For example, the sum of squares for A is computed as the difference between the sum of squares for the model A+C+AC and the sum of squares for the model C+AC. The sum of squares for C and AC is computed similarly.

Next, the sum of squares of the subject factor is computed by treating the subjects as a one-way design, computing the subject sum of squares, and subtracting the sum of squares of all terms that occur before it in the model—in this case, the sum of squares of factor A.

Next, the sum of squares of the BC(A) interaction is computed by treating this term as a one-way design, computing its sum of squares, and subtracting the sum of squares of all terms that occur before it in the model—in this case, the sum of squares for A, B(A), C, and AC.

The computations are carried out in this manner to give reasonable tests in the cases when there are unequal numbers of subjects per group or some subjects have missing measurements. The results are similar to the Type III sum of squares computations given by SAS.

### Mean Square

An estimate of the variation accounted for by this term. It is the sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and the mean square of its corresponding error term. This is also called the F-test value.

### Prob Level

The significance level of the above F-ratio, or the probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant. Note that if the value is significant at the specified value of alpha, a star is placed to the right of the F-Ratio.

This F-ratio is only valid if all the assumptions are valid. You should study the results of the preliminary tests to determine if the assumptions hold.

### Power (Alpha=0.05)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of type II error (β). The power of the test depends on the sample size, the magnitudes of the variances, the alpha level, and the actual differences among the population means.

The power value calculated here assumes that the population standard deviation is equal to the observed standard deviation and that the differences among the population means are exactly equal to the differences among the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false. This is a critical measure of precision in hypothesis testing.

Generally, you would consider the power of the test when you accept the null hypothesis. The power will give you some idea of what actions you might take to make your results significant. If you accept the null hypothesis with high power, there is not much left to do. At least you know that the means are not different. However, if you accept the null hypothesis with low power, you can take one or more of the following actions:

1. Increase your alpha level. Perhaps you should be testing at alpha = 0.05 instead of alpha = 0.01. Increasing the alpha level will increase the power.
2. Increasing your sample size will increase the power of your test if you have low power. If you have high power, an increase in sample size will have little effect.
3. Decrease the magnitude of the variance. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

## Probability Levels for F-Tests with Geisser-Greenhouse Adjustments

**Probability Levels for F-Tests with Geisser-Greenhouse Adjustments**

| Source Term | DF | F-Ratio | Regular Prob Level | Lower Bound Epsilon Prob Level | Geisser Greenhouse Epsilon Prob Level | Huynh Feldt Epsilon Prob Level |
|---|---|---|---|---|---|---|
| A: Exercise | 2 | 0.61 | 0.555040 | | | |
| B(A): Subject | 15 | | | | | |
| C: Time | 2 | 36.92 | 0.000000* | 0.000021* | 0.000000* | 0.000000* |
| AC | 4 | 6.45 | 0.000716* | 0.009496* | 0.000755* | 0.000716* |
| BC(A) | 30 | | | | | |
| S | 0 | | | | | |

This table presents the F ratios from the analysis of variance table with probability levels computed using the three Geisser-Greenhouse adjustments. These are explained in detail below.

Note that no adjustments are made to between-subjects terms (A in this example). Also note that in designs involving two or three within factors, different adjustment factors are computed for each term. The values of epsilon are shown in the Covariance Matrix Circularity report.

### Source Term

The source of variation, which is the term in the model.

### F-Ratio

The F-ratio is repeated from Analysis of Variance Table.

### Regular Prob Level

The probability level is repeated from Analysis of Variance Table.

### Lower-Bound Epsilon Prob Level

This is the probability level of the corresponding F-ratio using the minimum epsilon. This correction involves multiplying both the numerator and denominator degrees of freedom by the minimum epsilon and then calculating the probability level. Since this epsilon is a value between zero and one, the impact of this adjustment is to reduce the degrees of freedom.

This adjustment is made to correct for a non-circular covariance matrix. Simulation studies have shown these probability levels to be too conservative and so we do not recommend its use. Usually, the Geisser-Greenhouse epsilon is used instead.

### Geisser-Greenhouse Epsilon Prob Level

This is the probability level of the corresponding F-ratio using the Geisser-Greenhouse epsilon. This adjustment involves multiplying both the numerator and denominator degrees of freedom by the Geisser-Greenhouse epsilon and then calculating the probability level. Since this epsilon is a value between zero and one, the impact of this adjustment is to reduce the degrees of freedom.

This adjustment is made to correct for non-circularity in the covariance matrix. Box suggested that rather than using the theoretical minimum value of the Geisser-Greenhouse epsilon, you should use the value estimated by the data.

Simulation studies have shown this adjustment to give very accurate probability levels. We recommend its use.

### Huynh-Feldt Epsilon Prob Level

This is the probability level of the corresponding F-ratio using the Huynh- Feldt version of the Geisser-Greenhouse correction. This correction involves multiplying both the numerator and denominator degrees of freedom by their epsilon and then calculating the probability level. Since this epsilon is a value between zero and one, the impact of this adjustment is to reduce the degrees of freedom.

This adjustment is made to correct for non-circularity in the covariance matrix. Huynh and Feldt showed that Geisser-Greenhouse estimate of epsilon was biased so they developed a less biased version. When this estimate is greater than one, it is set equal to one.

Simulation studies have shown this adjustment to give accurate probability levels, but not as accurate as Geisser-Greenhouse correction. Hence, we recommend the Geisser-Greenhouse correction.

## Strategy for the Geisser-Greenhouse Adjustment

Kirk (1982) recommends the following three step testing strategy.

1. Check the Regular Prob Level. If this probability level is not significant (if it is not lower than 0.05, say), stop and declare the F not significant. If this F is significant, proceed to step 2.
2. Check the Lower-Bound Prob Level. If this probability is significant (less than 0.05, say), stop and declare the F significant. If this F is not significant, proceed to step 3.
3. Check the Geisser-Greenhouse Prob Level. If this probability is significant, stop and declare the F significant. If this probability level is not significant, declare the F as not significant.

## Power Values for F-Tests with Geisser-Greenhouse Adjustments

**Power Values for F-Tests with Geisser-Greenhouse Adjustments**

| Source Term | DF | F-Ratio | Regular Power (Alpha=0.05) | Lower Bound Epsilon Power (Alpha=0.05) | Geisser Greenhouse Epsilon Power (Alpha=0.05) | Huynh Feldt Epsilon Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Exercise | 2 | 0.61 | 0.133339 | | | |
| B(A): Subject | 15 | | | | | |
| C: Time | 2 | 36.92 | 1.000000 | 0.999890 | 1.000000 | 1.000000 |
| AC | 4 | 6.45 | 0.977632 | 0.834571 | 0.976634 | 0.977632 |
| BC(A) | 30 | | | | | |
| S | 0 | | | | | |

This table presents the F ratios from the analysis of variance table with the associated power values. The definition of power is discussed above in the Analysis of Variance section. This table lets you compare the statistical power of the four tests.

Note how the power decreases as the more conservative tests are used. Since the Geisser-Greenhouse is the most conservative test, it has the lowest power.

### Source Term

The source of variation, which is the term in the model.

### F-Ratio

The F-ratio is repeated from Analysis of Variance Table.

### Regular Power (Alpha=0.05

This gives the power. The definition of power and how it is calculating was provided in the Analysis of Variance Table section.

### Lower-Bound Epsilon Power (Alpha=0.05)

This gives the power when the Lower-Bound correction is used.

### Geisser-Greenhouse Power Epsilon (Alpha=0.05)

This gives the power when the Geisser-Greenhouse correction is used.

### Huynh Feldt Epsilon Power (Alpha=0.05)

This gives the power when the Huynh Feldt correction is used.

# Box's M Test for Equality of Between-Group Covariance Matrices Section

**Box's M Test for Equality of Between-Group Covariance Matrices**

| Source Term | Box's M | DF1 | DF2 | F Value | Prob Level | Chi2 Value | Prob Level | Covariance Matrices Equal? |
|---|---|---|---|---|---|---|---|---|
| BC(A) | 16.94 | 12.0 | 1090.4 | 0.99 | 0.457845 | 12.05 | 0.441998 | Okay |

This section presents the results of a preliminary test to determine if the data meet the assumption of equal covariance matrices across groups. This test is discussed in detail in the Equality of Covariance Matrices chapter. Since the test depends heavily on the assumption of multivariate normality, when the data fail to pass the test, it may or may not be because of the covariances matrices are unequal.

When your data fail this test, one remedy is to transform the response variable by taking the square root, the logarithm, or the inverse. Often, a power transformation such as these will correct both non-normality and unequal variance. Of course, after applying such a variance stabilizing transformation, you have to discuss your results in the transformed metric—you cannot discussed the means in the original (untransformed) metric.

Note that this test requires the number of subjects per group to be greater than the number of levels of the within-subject factor(s).

### Source Term

This is the term whose covariance matrices are being tested. The factor in parentheses represents the term(s) forming the groups, the first factor listed (B in this example) is the subject factor, and the rest of the factors are used to form the multivariate response. In this example, factor C, which has three levels, becomes the multivariate response vector. If more than one factor is listed here, they are combined into one single factor to form the multivariate response vector.

### Box’s M

This is the value of Box’s M statistic used to test the equality of the covariance matrices.

### DF1

The numerator degrees of freedom of the approximate F-distribution used to evaluate Box’s M statistic. This value need not be an integer. This value is also the degrees of freedom of the approximate Chi-square statistic.

### DF2

The denominator degrees of freedom of the approximate F-distribution used to evaluate Box’s M statistic. This value need not be an integer.

### F Value

The value of the approximate F-test used to evaluate Box’s M statistic.

### Prob Level

The probability level of the Box’s M statistic based on an F-distribution approximation. If this value is less than a chosen significance value, such as 0.10, you must assume that the covariance matrices are not equal and take appropriate action.

### Chi2 Value

The value of the approximate Chi-square test used to evaluate Box’s M statistic.

### Prob Level

The probability level of the Box's M statistic based on a Chi-square approximation. If this value is less than a chosen significance value, such as 0.10, you must assume that the covariance matrices are not equal and take appropriate action.

### Covariance Matrices Equal?

Using the value of the Assumption Alpha contained on the Reports tab panel, this provides the result of the test.

## Covariance Matrix Circularity Section

**Covariance Matrix Circularity Section**

| Source Term | Lower Bound Epsilon | Geisser Greenhouse Epsilon | Huynh Feldt Epsilon | Mauchly Test Statistic | Chi2 Value | DF | Prob Level | Covariance Matrix Circularity? |
|---|---|---|---|---|---|---|---|---|
| BC(A) | 0.500000 | 0.989629 | 1.000000 | 0.989521 | 0.1 | 2.0 | 0.928911 | Okay |

Note: Mauchly's statistic actually tests the more restrictive assumption that the pooled covariance matrix has compound symmetry.

This section provides an analysis of the circularity (sometimes called the sphericity) assumption that is required for all of the within-subject F tests. The formulas are given in the Technical Details at the beginning of the chapter. You can often correct circularity problems by taking the logarithm of the responses.

Some statisticians believe you should ignore this test since it relies heavily on the multivariate normality of your data. They suggest that you routinely use Box's Geisser-Greenhouse correction which corrects for this problem.

### Source Term

This is the term whose covariance matrix is being tested for circularity. The factor in parentheses represents the term(s) forming the groups, the first factor listed (B in this example) is the subject factor, and the rest of the factors are used to form the multivariate response. In this example, factor C, which has three levels, becomes the multivariate response vector. If more than one factor is listed, they are combined into one single factor to form the multivariate response vector.

### Lower Bound Epsilon

This is the minimum value of epsilon. The maximum value is one. This value is used to adjust the F-test by multiplying it times both the numerator and denominator degrees of freedom when calculating the probability levels.

### Geisser Greenhouse Epsilon

This is the estimate of epsilon that was suggested by Box. It serves as an index of the severity of non-circularity. Values of epsilon near one indicate that the covariance matrix is circular. Values of epsilon near the minimum (the Lower Bound Epsilon) indicate that the covariance matrix assumption is violated.

This value is used to adjust the F-test by multiplying it times both the numerator and denominator degrees of freedom when calculating the probability levels.

### Huynh Feldt Epsilon

This is an estimate of epsilon that was suggested by Huynh and Feldt to correct for bias found it the Geisser Greenhouse estimate. This estimate is always greater than or equal to the Geisser-

Greenhouse estimate. It is possible for this value to be greater than one. When this happens, the value is set equal to one.

Epsilon serves as an index of the severity of non-circularity. Values near one indicate that the covariance matrix is circular. Values near the minimum (the Lower Bound Epsilon) indicate that the covariance matrix assumption is violated.

This value is used to adjust the F-test by multiplying it times both the numerator and denominator degrees of freedom when calculating the probability levels.

### Mauchly Test Statistic

This is the value of Mauchly's test statistic. It tests the assumption that the pooled covariance matrix has compound symmetry. Compound symmetry is slightly more restrictive than circularity. The value of this statistic ranges from zero to one.

### Chi2 Value

This chi-square value is used to test the significance of the Mauchly test statistic.

### DF

This is the degrees of freedom of the chi-square approximation of Mauchly's test statistic.

### Prob Level

This is the significance level of the chi-square test. When this value is small (0.10 or less), the data fail Mauchly's test for compound symmetry.

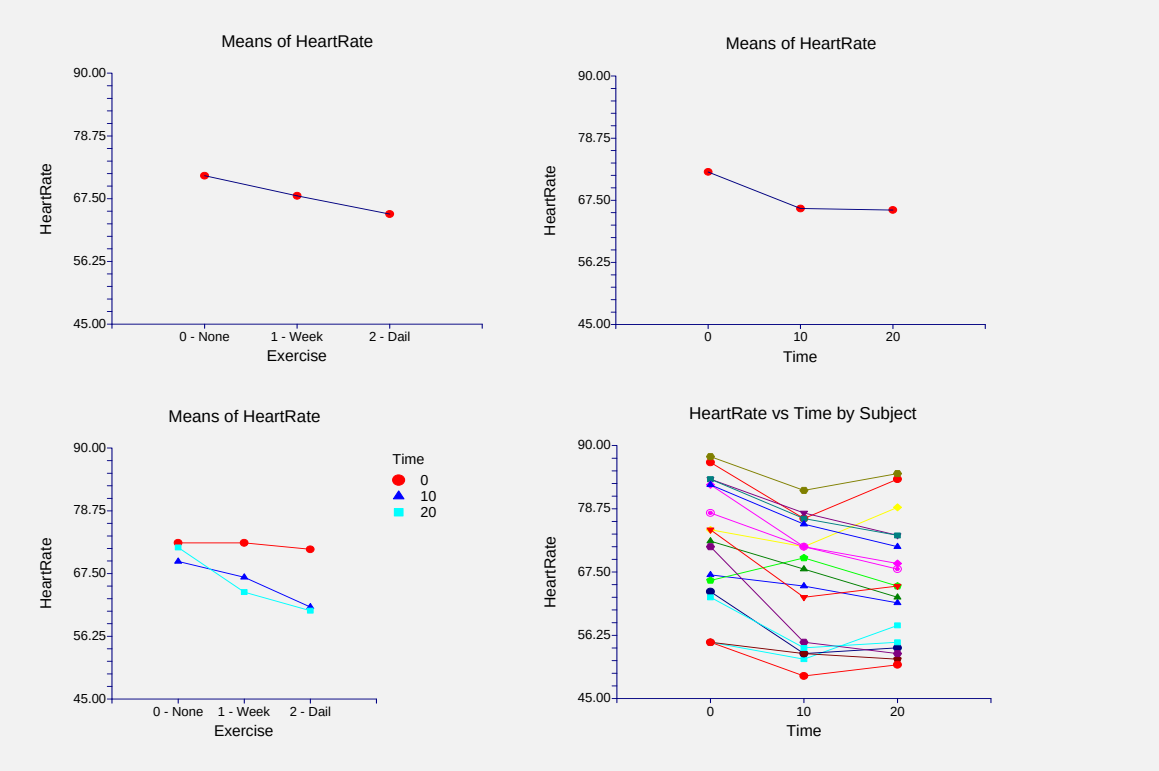
### Covariance Matrix Circularity?

This field indicates whether the data passed or failed Mauchly's test.

## Means, Standard Errors, and Plots Sections

**Means and Standard Errors Section**

| Term | Count | Mean | Standard Error |
|---|---|---|---|
| All | 54 | 68.11111 | |
| A: Exercise | | | |
| 0 - None | 18 | 71.61111 | 4.403095 |
| 1 - Weekly | 18 | 68 | 4.403095 |
| 2 - Daily | 18 | 64.72222 | 4.403095 |
| C: Time | | | |
| 0 | 18 | 72.61111 | 0.641821 |
| 10 | 18 | 66 | 0.641821 |
| 20 | 18 | 65.72222 | 0.641821 |
| AC: Exercise,Time | | | |
| 0 - None,0 | 6 | 73 | 1.111667 |
| 0 - None,10 | 6 | 69.66666 | 1.111667 |
| 0 - None,20 | 6 | 72.16666 | 1.111667 |
| 1 - Weekly,0 | 6 | 73 | 1.111667 |
| 1 - Weekly,10 | 6 | 66.83334 | 1.111667 |
| 1 - Weekly,20 | 6 | 64.16666 | 1.111667 |
| 2 - Daily,0 | 6 | 71.83334 | 1.111667 |
| 2 - Daily,10 | 6 | 61.5 | 1.111667 |
| 2 - Daily,20 | 6 | 60.83333 | 1.111667 |




### Term

The label for this line of the report.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Standard Error

The standard error of the mean. Note that these standard errors are the square root of the mean square of the error term for this term divided by the count. These standard errors are not the same as the simple standard errors calculated separately for each group. The standard errors reported here are those appropriate for testing multiple comparisons.

### Plot of Means

These plots display the means for each factor and two-way interaction. Note how easily you can see patterns in the plots.

## Multiple Comparison Section

**Tukey-Kramer Multiple-Comparison Test**

Response: HeartRate
Term A: Exercise

Alpha=0.050 Error Term=B(A) DF=15 MSE=348.9704 Critical Value=3.673397

| Group | Count | Mean | Different From Groups |
|---|---|---|---|
| 2 - Daily | 18 | 64.72222 | |
| 1 - Weekly | 18 | 68 | |
| 0 - None | 18 | 71.61111 | |

**Tukey-Kramer Multiple-Comparison Test**

Response: HeartRate
Term C: Time

Alpha=0.050 Error Term=BC(A) DF=30 MSE=7.414815 Critical Value=3.486436

| Group | Count | Mean | Different From Groups |
|---|---|---|---|
| 20 | 18 | 65.72222 | 0 |
| 10 | 18 | 66 | 0 |
| 0 | 18 | 72.61111 | 20, 10 |

These sections present the results of the multiple-comparison procedures selected. These reports all use a uniform format that will be described by considering Tukey-Kramer Multiple-Comparison Test. The reports for the other procedures are similar. For more information on the interpretation of the various multiple-comparison procedures, turn to the section by that name in the One-Way ANOVA chapter.

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom for the error term.

### MSE

The value of the mean square error.

### Critical Value

The value of the test statistic that is “just significant” at the given value of alpha. This value depends on which multiple-comparison procedure you are using. It is based on the t-distribution or the studentized range distribution. It is the value of t, F, or q in the corresponding formulas.

### Group

The label for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

### Different from Groups

A list of those groups that are significantly different from this group according to this multiple-comparison procedure. All groups not listed are not significantly different from this group.

## Planned Comparison Section

This section presents the results of any planned comparisons that were selected.

**Planned Comparison: A: 0 - None vs. 1 - Weekly**

Response: HeartRate
Term A: Exercise

Alpha=0.050 Error Term=B(A) DF=15 MSE=348.9704

Comparison Value=-3.611111 T-Value=0.5799196 Prob>|T|=0.570578 Decision(0.05)=Do Not Reject
Comparison Standard Error=6.226916

| Group | Comparison Coefficient | Count | Mean |
|---|---|---|---|
| 0 - None | -1 | 18 | 71.61111 |
| 1 - Weekly | 1 | 18 | 68 |
| 2 - Daily | 0 | 18 | 64.72222 |

**Planned Comparison: 0 - None vs. 2 - Daily**

Response: HeartRate
Term A: Exercise

Alpha=0.050 Error Term=B(A) DF=15 MSE=348.9704

Comparison Value=-6.888889 T-Value=1.106308 Prob>|T|=0.286021 Decision((0.05)=Do Not Reject
Comparison Standard Error=6.226916

| Group | Comparison Coefficient | Count | Mean |
|---|---|---|---|
| 0 - None | -1 | 18 | 71.61111 |
| 1 - Weekly | 0 | 18 | 68 |
| 2 - Daily | 1 | 18 | 64.72222 |

### Alpha

The level of significance that you selected.

### Error Term

The term in the ANOVA model that is used as the error term.

### DF

The degrees of freedom of the error term.

### MSE

The value of the mean square error.

### Comparison Value

The value of the comparison. This is formed by multiplying the Comparison Coefficient times the Mean for each group and summing.

### T-Value

The t-test used to test whether the above Comparison Value is significantly different from zero.

$$t_f = \frac{\sum_{i=1}^{k} c_i M_i}{\sqrt{MSE \sum_{i=1}^{k} \frac{c_i^2}{n_i}}}$$

where ***MSE*** is the mean square error, $f$ is the degrees of freedom associated with ***MSE***, ***k*** is the number of groups, $c_i$ is the comparison coefficient for the $i^{th}$ group, $M_i$ is the mean of the $i^{th}$ group, and $n_i$ is the sample size of the $i^{th}$ group.

### Prob>|T|

The significance level of the above T-Value. The Comparison is statistically significant if this value is less than the specified alpha.

### Decision(0.05)

The decision based on the specified value of the multiple comparison alpha.

### Comparison Standard Error

This is the standard error of the estimated comparison value. It is the denominator of the T-Value (above).

### Group

The label for this group.

### Comparison Coefficient

The coefficient (weight) used for this group.

### Count

The number of observations in the mean.

### Mean

The value of the sample mean.

# Example 2 – Single-Group Repeated-Measures Design

This section presents an example of how to analyze a single-group repeated measures design. The dataset was given at the beginning of the chapter and is contained in the REACTION database.

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Repeated Measures Analysis of Variance window.

**1 Open the REACTION dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **REACTION.S0**.
- Click **Open**.

**2 Open the Repeated Measures Analysis of Variance window.**

- On the menus, select **Analysis**, then **Analysis of Variance (ANOVA)**, then **Repeated Measures Analysis of Variance**. The Repeated Measures Analysis of Variance procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Repeated Measures Analysis of Variance window, select the **Variables tab**.
- Double-click in the **Response Variable(s)** box. This will bring up the variable selection window.
- Select **Test** from the list of variables and then click **Ok**.
- Double-click in the **Subject Variable** box. This will bring up the variable selection window.
- Select **Person** from the list of variables and then click **Ok**.
- Clear the value in the **Between Factor 1** box.
- Double-click in the **Within Factor 1** box. This will bring up the variable selection window.
- Select **Drug** from the list of variables and then click **Ok**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The output will appear as follows:

# Single-Group Repeated Measures Output

**Repeated Measures ANOVA Report**

Database Reaction.S0
Response Test

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A: Person | 4 | No | S(AB) | S+bsA |
| B: Drug | 3 | Yes | AB | S+sAB+asB |
| AB | 12 | No | S(AB) | S+sAB |
| S(AB) | 0 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Person | 4 | 680.8 | 170.2 | | | |
| B: Drug | 3 | 698.2 | 232.7333 | 24.76 | 0.000020* | 0.999998 |
| AB | 12 | 112.8 | 9.4 | | | |
| S | 0 | | | | | |
| Total (Adjusted) | 19 | 1491.8 | | | | |
| Total | 20 | | | | | |

* Term significant at alpha = 0.05

**Probability Levels for F-Tests with Geisser-Greenhouse Adjustments**

| Source Term | DF | F-Ratio | Regular Prob Level | Lower Bound Epsilon Prob Level | Geisser Greenhouse Epsilon Prob Level | Huynh Feldt Epsilon Prob Level |
|---|---|---|---|---|---|---|
| A: Person | 4 | | | | | |
| B: Drug | 3 | 24.76 | 0.000020* | 0.007620* | 0.000649* | 0.000020* |
| AB | 12 | | | | | |
| S | 0 | | | | | |

**Power Values for F-Tests with Geisser-Greenhouse Adjustments Section**

| Source Term | DF | F-Ratio | Regular Power (Alpha=0.05) | Lower Bound Epsilon Power (Alpha=0.05) | Geisser Greenhouse Epsilon Power (Alpha=0.05) | Huynh Feldt Epsilon Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Person | 4 | | | | | |
| B: Drug | 3 | 24.76 | 0.999998 | 0.953259 | 0.998877 | 0.999998 |
| AB | 12 | | | | | |
| S | 0 | | | | | |

**Covariance Matrix Circularity Section**

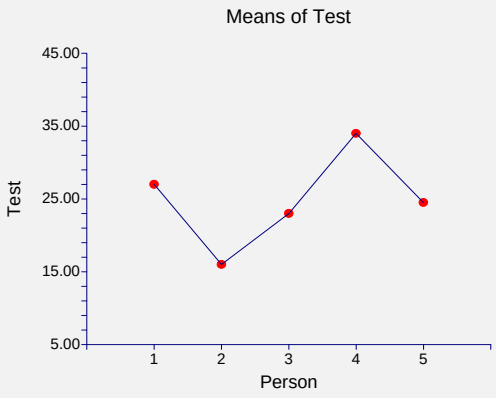
| Source Term | Lower Bound Epsilon | Geisser Greenhouse Epsilon | Huynh Feldt Epsilon | Mauchly Test Statistic | Chi2 Value | DF | Prob Level | Covariance Matrix Circularity? |
|---|---|---|---|---|---|---|---|---|
| AB | 0.333333 | 0.604874 | 1.000000 | 0.186495 | 4.6 | 5.0 | 0.470366 | Okay |

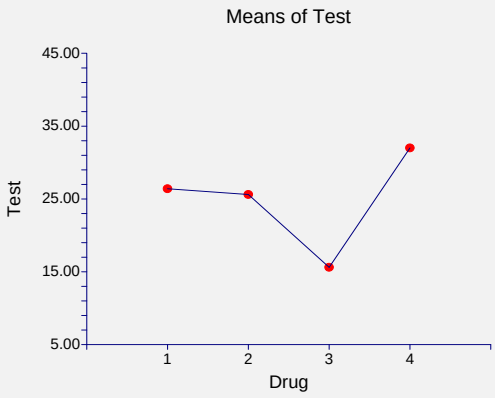
Note: Mauchly's statistic actually tests the more restrictive assumption that the pooled covariance matrix has compound symmetry.

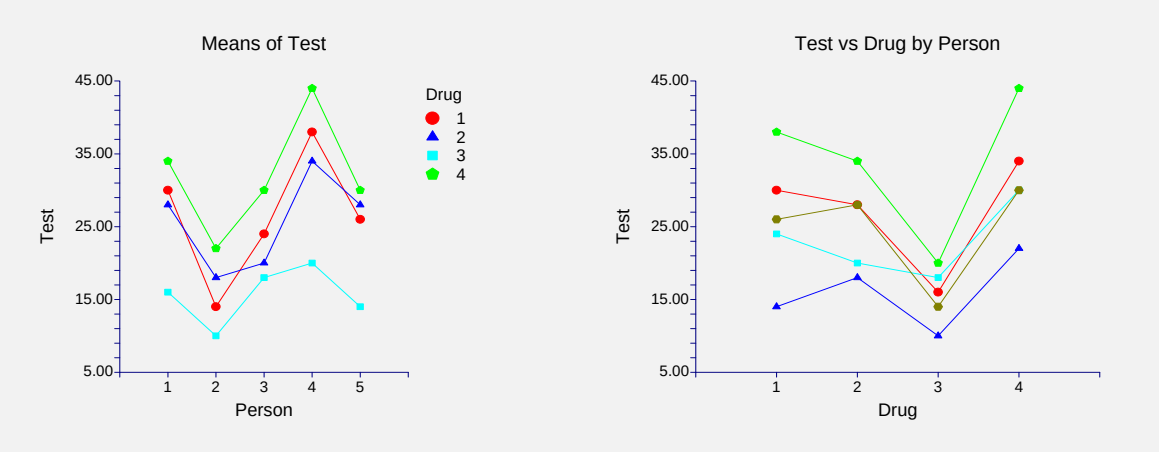
### Means and Standard Error Section

| Term | Count | Mean | Standard Error |
|---|---|---|---|
| All | 20 | 24.9 | |
| A: Person | | | |
| 1 | 4 | 27 | 0 |
| 2 | 4 | 16 | 0 |
| 3 | 4 | 23 | 0 |
| 4 | 4 | 34 | 0 |
| 5 | 4 | 24.5 | 0 |
| | | | |
| B: Drug | | | |
| 1 | 5 | 26.4 | 1.371131 |
| 2 | 5 | 25.6 | 1.371131 |
| 3 | 5 | 15.6 | 1.371131 |
| 4 | 5 | 32 | 1.371131 |
| AB: Person,Drug | | | |
| 1,1 | 1 | 30 | 0 |
| 1,2 | 1 | 28 | 0 |
| 1,3 | 1 | 16 | 0 |
| 1,4 | 1 | 34 | 0 |
| 2,1 | 1 | 14 | 0 |
| 2,2 | 1 | 18 | 0 |
| 2,3 | 1 | 10 | 0 |
| 2,4 | 1 | 22 | 0 |
| 3,1 | 1 | 24 | 0 |
| 3,2 | 1 | 20 | 0 |
| 3,3 | 1 | 18 | 0 |
| 3,4 | 1 | 30 | 0 |
| 4,1 | 1 | 38 | 0 |
| 4,2 | 1 | 34 | 0 |
| 4,3 | 1 | 20 | 0 |
| 4,4 | 1 | 44 | 0 |
| 5,1 | 1 | 26 | 0 |
| 5,2 | 1 | 28 | 0 |
| 5,3 | 1 | 14 | 0 |
| 5,4 | 1 | 30 | 0 |

### Plots Section






Our only comment about this output is to note that the Box's M test section was omitted because there is only one group.

# Chapter 220

# Mixed Models

## Introduction

The Mixed Models procedure analyzes results from a wide variety of experimental designs in which the outcome (response) is continuous, including

- Two-sample designs (replacing the t-test)
- One-way layout designs (replacing one-way ANOVA)
- Factorial designs (replacing factorial GLM)
- Split-plot designs (replacing split-plot GLM)
- Repeated-measures designs (replacing repeated-measures GLM)
- Cross-over designs (replacing GLM)
- Designs with covariates (replacing GLM)

The Mixed Models procedure can be used to test and estimate means (including pair-wise comparisons among levels), compare models, estimate variance-covariance matrix components, and produce graphs of means and repeated measurements of subjects. Examples are given in this chapter of models with only between-subjects factors, only within-subjects factors, and both between- and within-subjects factors. Analysis of covariance examples and multiple comparisons examples are also included.

## Why Use a Mixed Model?

As stated above, mixed models have several advantages over traditional linear models. Just a few are listed here.

- **Specifying More Appropriate Variance-Covariance Structures for Longitudinal Data**: The ability to fit complex covariance patterns provides more appropriate fixed effect estimates and standard errors.
- **Analysis Assuming Unequal Group Variances**: Different variances can be fit for each treatment group.
- **Analysis of Longitudinal Data with Unequal Time Points**: Mixed models allow for the analysis of data in which the measurements were made at random (varying) time points.
- **Analysis of Longitudinal Data with Missing Response Data**: Problems caused by missing data in repeated measures and cross-over trials are eliminated.
- **Greater Flexibility in Modeling Covariates**: Covariates can be modeled as fixed or random and more accurately represent their true contribution in the model.

Mixed models are particularly useful in medical studies where a wide variety of factors influence the response to a treatment of interest. For example, suppose that an experimental treatment is being administered to a group of patients desiring to lose weight. Traditional statistical methodologies (e.g., ANOVA, multiple regression, etc.) require that the treatments be given at

the same time intervals for all patients in the group in order for the statistical analysis and conclusions to be accurate. What would happen if patients were not all able to receive the treatment at the same time intervals or if some patients missed some treatments? Traditional statistical approaches would no longer be valid since there are random events or components entering into the experiment. This is where mixed models techniques become useful. A mixed model would allow us to make inferences about the treatment by modeling and estimating the random components. Furthermore, mixed models allow us to make greater use of incomplete data, such as that obtained from patients who drop out or miss scheduled treatments. Traditional methods would exclude such individuals from the analysis, losing valuable information.

## What is a Mixed Model?

In a general linear model (GLM), a random sample of the individuals in each population is drawn. A treatment is applied to each individual in the sample and an outcome is measured. The data so obtained are analyzed using an analysis of variance table that produces an F-test.

A mathematical model may be formulated that underlies each analysis of variance. This model expresses the response variable as the sum of parameters of the population. For example, a linear model for a two-factor experiment could be

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$$

where $i = 1, 2, \ldots, I$ (the number of levels of factor 1), $j = 1, 2, \ldots, J$ (the number of levels of factor 2), and $k = 1, 2, \ldots, K$ (the number of subjects in the study). This model expresses the value of the response variable, $Y$, as the sum of five components:

- $\mu$ the mean.
- $a_i$ the contribution of the $i^{th}$ level of a factor A.
- $b_j$ the contribution of the $j^{th}$ level of a factor B.
- $(ab)_{ij}$ the combined contribution (or interaction) of the $i^{th}$ level of a factor A and the $j^{th}$ level of a factor B.
- $e_{ijk}$ the contribution of the $k^{th}$ individual. This is often called the “error.”

In this example, the linear model is made up of *fixed effects* only. An effect is fixed if the levels in the study represent all levels of the factor that are of interest, or at least all levels that are important for inference (e.g., treatment, dose, etc.).

The following assumptions are made when using the F-test in a general linear model.

1. The response variable is continuous.
2. The individuals are independent.
3. The $e_{ijk}$ follow the normal probability distribution with mean equal to zero.
4. The variances of the $e_{ijk}$ are equal for all values of $i$, $j$, and $k$.

The *Linear Mixed Model* (or just *Mixed Model*) is a natural extension of the general linear model. Mixed models extend linear models by allowing for the addition of *random effects*, where the levels of the factor represent a random subset of a larger group of all possible levels (e.g., time of administration, clinic, etc.). For example, the two-factor linear model above could be augmented to include a random block effect such as clinic or doctor since the clinic or doctor may be assumed to be a random realization from a distribution of clinics or doctors. Covariates

(continuous) and/or nested effects can also be included in the mixed model to improve the accuracy of the fixed effect estimates. The general form of the mixed model in matrix notation is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$$

where

- $\mathbf{y}$ vector of responses
- $\mathbf{X}$ known design matrix of the fixed effects
- $\boldsymbol{\beta}$ unknown vector of fixed effects parameters to be estimated
- $\mathbf{Z}$ known design matrix of the random effects
- $\mathbf{u}$ unknown vector of random effects
- $\boldsymbol{\varepsilon}$ unobserved vector of random errors

We assume

$\mathbf{u} \sim N(\mathbf{0},\mathbf{G})$

$\boldsymbol{\varepsilon} \sim N(\mathbf{0},\mathbf{R})$

$\text{Cov}[\mathbf{u}, \boldsymbol{\varepsilon}] = \mathbf{0}$

where

- $\mathbf{G}$ variance-covariance matrix of $\mathbf{u}$
- $\mathbf{R}$ variance-covariance matrix of the errors $\boldsymbol{\varepsilon}$

The variance of $\mathbf{y}$, denoted $\mathbf{V}$, is

$$\begin{aligned}\mathbf{V} &= \text{Var}[\mathbf{y}] \\ &= \text{Var}[\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}] \\ &= 0 + \text{Var}[\mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}] \\ &= \mathbf{ZGZ'} + \mathbf{R}\end{aligned}$$

In order to test the parameters in $\boldsymbol{\beta}$, which is typically the goal in mixed model analysis, the unknown parameters ($\boldsymbol{\beta}$, $\mathbf{G}$, and $\mathbf{R}$) must be estimated. Estimates for $\boldsymbol{\beta}$ require estimates of $\mathbf{G}$ and $\mathbf{R}$. In order to estimate $\mathbf{G}$ and $\mathbf{R}$, the structure of $\mathbf{G}$ and $\mathbf{R}$ must be specified. Details of the specific structures for $\mathbf{G}$ and $\mathbf{R}$ are discussed later.

The following assumptions are made when using the F-test in a mixed model.

1. The response variable is continuous.
2. The individuals are independent.
3. The random error follows the normal probability distribution with mean equal to zero.

A distinct (and arguably the most important) advantage of the mixed model over the general linear model is flexibility in random error and random effect variance component modeling (note that the equal-variance assumption of the general linear model is not necessary for the linear mixed model). Mixed models allow you to model both heterogeneous variances and correlation among observations through the specification of the covariance matrix structures for $\mathbf{u}$ and $\boldsymbol{\varepsilon}$. You should be careful to build an appropriate covariance structure for the model, since the hypothesis tests, confidence intervals, and treatment mean estimates are all affected by the covariance structure of the model. The variance matrix estimates are obtained using maximum likelihood

(ML) or, more commonly, restricted maximum likelihood (REML). The fixed effects in the mixed model are tested using F-tests. More details of random factor estimation and fixed factor estimation and testing are given later in this chapter.

## Types of Mixed Models

Several general mixed model subtypes exist that are characterized by the random effects, fixed effects, covariate terms, and covariance structure they involve. These include fixed effects models, random effects models, covariance pattern models, and random coefficients models.

### Fixed Effects Models

A *fixed effects model* is a model where only fixed effects are included in the model. An effect (or factor) is fixed if the levels in the study represent all levels of interest of the factor, or at least all levels that are important for inference (e.g., treatment, dose, etc.). No random components are present. The general linear model is a fixed effects model. Fixed effects models can include covariates and/or interactions. The two-factor experiment example above gives an example of a fixed effects model. The fixed effects can be estimated and tested using the F-test. Fixed effects are specified as the Fixed Factors Model on the Variables tab.

Note: If only one response is recorded for each subject, there is no within-subject correlation to be modeled in variance-covariance matrix. If more than one response is measured for each subject, you could use repeated measures ANOVA or use a random-coefficients mixed model.

### Random Effects Models

A *random effects model* is a model with only random terms in the model. An effect (or factor) is random if the levels of the factor represent a random subset of a larger group of all possible levels (e.g., patients represent the population as a whole). Random effects are specified in the Subject (Random) Model box on the Variables tab. The random effects are not tested, but estimates are given.

Note: If only one response is recorded for each subject, there is no within-subject correlation to be modeled in variance-covariance matrix. If more than one response is measured for each subject, you could use repeated measures ANOVA or use a random coefficients mixed model.

### Longitudinal Data Models

Longitudinal data arises when more than one response is measured on each subject in the study. Responses are often measured over time at fixed or random intervals. An interval is fixed if the measurements are made a pre-specified time intervals, e.g. measuring heart rate after 2 hours, 4 hours, and 6 hours after drug administration. An interval is random if the response measurements are made at random time points, e.g. measuring heart rate at the start of a race and after each runner finishes (presumably at differing time points). Various covariance structures can be employed to model the variance and correlation among repeated measurements or the relationship with time can be investigated. The manner in which the longitudinal data is modeled gives rise to two different mixed model subtypes: covariance pattern models and random coefficients models.

#### Covariance Pattern Models

If the covariance and correlation between repeated measurements is taken into account (i.e. modeled), the model is called a *covariance pattern model*. The covariance pattern model is usually appropriate if the repeated measurements occur at fixed intervals and the relationship with

time in not of particular interest. More information is given later in the chapter about the different covariance patterns that can be fit.

The repeated or residual covariance pattern is specified in the Repeated Variance Pattern box on the Variables tab.

### Random Coefficients Models

It is often important in a study to determine the relationship between the response and time. This is often done by including the measurement time as a covariate in the model, with a corresponding slope, say $\beta_t$. It is plausible and likely that the slope will vary with subject, so it might be useful to model a separate intercept and slope for each subject in the study. This is done by fitting the subject variable as the intercept and the subject*time interaction as the slope for each patient. These two terms could reasonably be assumed to arise at random from a distribution and, thus, would be specified as random effects. This gives rise to what is called a *random coefficients model*.

A random coefficients model is one in which the subject term and a subject*time interaction term are both included as random effects in the model. This type of model is different from an ordinary random effects model because when we fit a straight line, the estimates of the slope and intercept are not independent. Thus, the subject and subject*time effects in the model are correlated. The random effects model must be adapted to this situation to allow for correlation among these random effects. This is done using the bivariate normal distribution. The bivariate random effect becomes

$$\begin{pmatrix} subject_k \\ (subject * time)_k \end{pmatrix} \sim N(0, \mathbf{G}),$$

where

$$\mathbf{G} = \begin{pmatrix} \sigma^2_{subject} & \sigma_{subject,subject*time} \\ \sigma_{subject,subject*time} & \sigma^2_{subject*time} \end{pmatrix}.$$

The random coefficients model is usually used if the relationship with time is of interest or if the repeated measurements do not occur at fixed intervals. Random coefficient effects are specified in the Random Factor box on the Variables tab. Other fixed and random effects, besides time, can also be specified in the random coefficients model.

## Examples

Because of the large number of options, attempting to enter the appropriate model in the Mixed Models procedure can be intimidating. A number of examples at the end of the chapter are provided with the hope that one of the examples is similar enough to your scenario that it will guide you in selecting the options that are appropriate. The examples can also serve as a tutorial, beginning with the simplest two-group modeling in Example 2 (Example 1 is used for annotation) and continuing into more complex models.

Several of the examples also provide comparisons to analyses using classical procedures. For example, Example 3 compares the classical one-way analysis using the One-Way ANOVA procedure to the equivalent analysis using the Mixed Models procedure.

The examples at the end of this chapter are categorized in two ways.

1. The number of between-subject and within-subject factors
2. The experimental design or analysis method used

A brief explanation of between-subject factors and within-subject factors precedes the table of examples.

## Between-Subject Factors

Between-subject factors are those factors for which several subjects are assigned to (or sampled from) each level. If 12 subjects are randomly assigned to 3 treatment groups (4 subjects per group), treatment is a between-subject factor.

## Within-Subject Factors

Within-subject factors are those in which the subject's response is measure at several time points.

Within-subject factors are those factors for which multiple levels of the factor are measured on the same subject. If each subject is measured at the low, medium, and high level of the treatment, treatment is a within-subject factor.

## Example Overview

Example 1 has one within-subject factor and one between-subject factor, as well as a covariate. For Example 1, the output is annotated in detail. The remaining examples show the set-up and basic analysis.

| Example | Design/Analysis | Number of Between-Subject Factors | Number of Within-Subject Factors | Number of Covariates |
|---|---|---|---|---|
| 1 | Repeated Measures (+ Annotation) | 1 | 1 | 1 |
| 2a | Two-Group T-Test (Equal Variance) | 1 | 0 | 0 |
| 2b | Two-Group T-Test (Unequal Variance) | 1 | 0 | 0 |
| 2c | Two-Group T-Test (+ Covariate) | 1 | 0 | 1 |
| 3a | One-Way (Equal Variance) | 1 | 0 | 0 |
| 3b | One-Way (Unequal Variance) | 1 | 0 | 0 |
| 4 | One-Way (+ Covariate) | 1 | 0 | 1 |
| 5 | Factorial (+ Covariate) | 2 | 0 | 1 |
| 6 | RCBD | 0 | 1 | 0 |
| 7 | Complex Split-Plot | 1 | 2 | 2 |
| 8 | Cross-Over | 0 | 2 | 1 |
| 9 | Repeated Measures (Unequal Time Points) | 1 | 0 | 1 |

# Random versus Repeated Error Formulation

The general form of the linear mixed model as described earlier is

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}$$

$$\mathbf{u} \sim N(\mathbf{0},\mathbf{G})$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0},\mathbf{R})$$

$$\text{Cov}[\mathbf{u}, \boldsymbol{\varepsilon}] = \mathbf{0}$$

$$\mathbf{V} = \mathbf{ZGZ'} + \mathbf{R}$$

The specification of the random component of the model specifies the structure of **Z**, **u**, and **G**. The specification of the repeated (error or residual) component of the model specifies the structure of **ε** and **R**. Except in very complicated designs, it is recommended that only one of the two components be specified. That is, if the random component includes one or more terms, the repeated pattern should be the diagonal (basic) pattern. If the repeated pattern is more complicated than a diagonal, there should not be a random component. There are exceptions, but the resulting covariance structure should be carefully considered in such cases.

Specifying the random component of the model will suffice for most factorial, split-plot, and ANCOVA designs and for longitudinal designs with irregular time values. The repeated component of the model should be used for longitudinal analyses with a fixed number of time points (e.g., 1 hour, 2 hours, 4 hours, 8 hours), and where there are no, or very few, missing values.

In some scenarios, specifying a repeated pattern results in the same covariance parameter formulation as a random component. For example, specifying compound symmetry for the repeated pattern with no random component will result in the sample within-subject variance matrix as specifying Subject as the random factor and Diagonal for the repeated pattern. The examples of this chapter can be used to see the random and repeated specification for several common analyses.

# Determining the Correct Model of the Variance-Covariance of Y

## Akaike Information Criterion (AIC) for Model Assessment

Akaike information criterion (AIC) is tool for assessing model fit (Akaike, 1973, 1974). The formula is

$$AIC = -2 \times L + 2p$$

where $L$ is the (ML or REML) log-likelihood and $p$ depends on the type of likelihood selected. If the ML method is used, $p$ is the total number of parameters. If the REML method is used, $p$ is the number of variance component parameters.

The formula is designed so that a smaller AIC value indicates a “better” model. AIC penalizes models with larger numbers of parameters. That is, if a model with a much larger number of parameters produces only a slight improvement in likelihood, the values of AIC for the two models will suggest that the more parsimonious (limited) model is still the “better” model.

As an example, suppose a researcher would like to determine the appropriate variance-covariance structure for a longitudinal model with four equal time points. The researcher uses REML as the

likelihood type. The analysis is run five times, each with a different covariance pattern, and the AIC values are recorded as follows.

| Pattern | Number of Parameters | -2 log-likelihood | AIC |
|---|---|---|---|
| Diagonal | 1 | 214.43 | 216.43 |
| Compound Symmetry | 2 | 210.77 | 214.77 |
| AR(1) | 2 | 203.52 | 207.52 |
| Toeplitz | 4 | 198.03 | 206.03 |
| Unstructured | 7 | 197.94 | 211.94 |

The recommended variance-covariance structure among these five is the Toeplitz pattern, since it results in the smallest AIC value.

## What to Do When You Encounter a Variance Estimate that is Equal to Zero

It is possible that a mixed models data analysis results in a variance component estimate that is negative or equal to zero. This is particularly true in the case of random coefficients models. When this happens, the component that has a variance estimate equal to zero should be removed from the random factors model statement (or, if possible, the repeated pattern should be simplified to 'diagonal'), and the analysis should be rerun.

As an example, suppose a researcher would like to analyze a dataset using a random coefficients model. The data consists of sixty subjects, each of which received one of three treatments. The weight of each subject was measured at the beginning of the study and 6, 12 18, 24, and 30 days after administration of the treatment. The fixed and random factors models are entered as follows:

**Fixed Factors Model**: Day Trt Day*Trt

**Random Factors Model**: Subject Subject*Day

**Repeated (Time) Variance Pattern**: Diagonal

The mixed models analysis results in the following variance component parameter estimates:

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | **0.000000** | Subject |
| 1 | 2 | 0.031682 | Subject*Day |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 12.914745 | Diagonal (Variance) |

********** RUN ABORTED BECAUSE OF ZERO PARAMETER **********
Error Explanation:
One or more of the above parameter estimates is zero.
The corresponding term should not be included in the model.
The term must be removed from the model and then the problem rerun
in order to obtain the rest of the reports and charts.
**********

The estimated value for the Subject random component is equal to zero and should be removed from the analysis. Re-running the analysis without the Subject component in the random factors model results in the following parameter estimates:

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 0.030111 | Subject*Day |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 12.517215 | Diagonal (Variance) |

The variance estimates for the other parameters changed slightly after removing Subject from the random factors model.

## Fixed Effects

A fixed effect (or factor) is a variable for which levels in the study represent all levels of interest, or at least all levels that are important for inference (e.g., treatment, dose, etc.). The fixed effects in the model include those factor for which means, standard errors, and confidence intervals will be estimated and tests of hypotheses will be performed. Other variables for which the model is to be adjusted (that are not important for estimation or hypothesis testing) may also be included in the model as fixed factors. Fixed factors may be discrete variables or continuous covariates.

The correct model for fixed effects depends on the number of fixed factors, the questions to be answered by the analysis, and the amount of data available for the analysis. When more than one fixed factor may influence the response, it is common to include those factors in the model, along with their interactions (two-way, three-way, etc.). Difficulties arise when there are not sufficient data to model the higher-order interactions. In this case, some interactions must be omitted from the model. It is usually suggested that if you include an interaction in the model, you should also include the main effects (i.e. individual factors) involved in the interaction even if the hypothesis test for the main effects in not significant.

## Covariates

Covariates are continuous measurements that are not of primary interest in the study, but potentially have an influence on the response. Two types of covariates typically arise in mixed models designs: subject covariates and within-subject covariates. They are illustrated in the following example.

A study is conducted to determine the effect of two drugs on heart rate in mice. Each mouse receives each drug and a placebo with a washout period between treatments. The weight of each mouse is measured prior to the first treatment. The systolic blood pressure of each mouse is also measured immediately before each treatment. Although potentially an important factor, order of treatment is not considered in this example.

| Mouse | IWeight | Treatment | BP | HR |
|---|---|---|---|---|
| 1 | 18 | Placebo | 154 | 392 |
| 1 | 18 | Drug A | 167 | 378 |
| 1 | 18 | Drug B | 184 | 365 |
| 2 | 26 | Placebo | 166 | 402 |
| 2 | 26 | Drug A | 189 | 396 |
| 2 | 26 | Drug B | 177 | 397 |
| 3 | 22 | Placebo | 185 | 408 |
| 3 | 22 | Drug A | 163 | 402 |
| 3 | 22 | Drug B | 183 | 407 |
| 4 | 19 | Placebo | 167 | 411 |
| 4 | 19 | Drug A | 179 | 400 |
| 4 | 19 | Drug B | 172 | 392 |
| 5 | 15 | Placebo | 175 | 384 |
| 5 | 15 | Drug A | 168 | 391 |
| 5 | 15 | Drug B | 176 | 386 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |

In this example, initial weight (IWeight) and blood pressure (BP) are covariates. IWeight is a subject covariate because it is measured only once for each subject. BP is a within-subject covariate since it is measured on each subject for each treatment.

The Mixed Models procedure permits the user to make comparisons of fixed-effect means at specified values of covariates. For example, researchers could compare the two treatments to the placebo for IWeight = 20 and BP = 180, even when those values of the covariates do not appear in the actual data set.

Commonly, investigators wish to make comparisons of levels of a factor at several values of covariates. In this example, the researchers might want to compare the two treatments to the placebo at IWeight = 18, 23, and 26, and at BP = 160, 175, and 190. Caution should be exercised when making comparisons at multiple covariate values. The result in this case is $3 \times 3 = 9$ sets of comparisons and, therefore, $3 \times 9 = 27$ tests (3 pair wise treatment comparisons $\times$ 9 sets = 27 tests) for the Bonferroni adjustment of the p-value. After accounting for multiple testing, finding significant differences will require large sample sizes and/or extreme differences in means since the raw p-value would have to be less than 0.00185 in order to declare significance at the 0.05 level (0.05/27 = 0.00185).

# Time as a Fixed Effects Factor vs. Time as a Covariate

Time is an essential measurement in many mixed model designs. In some analyses, time may be considered a fixed factor, while in others it is covariate. A couple of examples illustrate this distinction.

## Time as a Fixed Effects Factor

Researchers wish to compare the extent to which rashes develop following administration of different doses of an anti-fungal cream. Fifteen individuals are divided into three groups, with each group receiving a different dose of the cream: low, medium, or high. The surface area of the resulting rash is measured at four time points: 1 hour, 2 hours, 4 hours, and 8 hours.

| Dose | Subject | Time | Rash |
|---|---|---|---|
| Low | 1 | 1 | 4.2 |
| Low | 1 | 2 | 3.5 |
| Low | 1 | 4 | 2.1 |
| Low | 1 | 8 | 6.8 |
| Low | 2 | 1 | 3.4 |
| Low | 2 | 2 | 5.2 |
| Low | 2 | 4 | 9.7 |
| Low | 2 | 8 | 6.5 |
| Low | 3 | 1 | 4.1 |
| Low | 3 | 2 | 6.8 |
| Low | 3 | 4 | 7.1 |
| Low | 3 | 8 | 2.3 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| High | 15 | 1 | 6.4 |
| High | 15 | 2 | 8.2 |
| High | 15 | 4 | 9.4 |
| High | 15 | 8 | 8.5 |

In this example, the time points are very structured (every subject is measured at the same time points) and the relationship between the size of the rash and time is not likely to be linear (the relationship will likely increase and then decrease). These two aspects of the study would generally lead the researcher to include Time as a fixed effects factor rather than as a covariate. If, however, the relationship were linear (or could be made linear by a suitable transformation), time could be considered a covariate. The next example examines the case where Time must be considered a covariate.

## Time as a Covariate

Three diets are compared for recently hatched chicks for their effect on growth. One hundred forty-seven chicks are randomly divided into three diets: low soybean protein, high soybean protein, and high fishmeal protein. Weights of chicks are measured at unequal times for two months after beginning the diet.

| Diet | Chick | Time | Weight |
|---|---|---|---|
| Low Soy | 1 | 5 | 64 |
| Low Soy | 1 | 11 | 69 |
| Low Soy | 1 | 24 | 74 |
| Low Soy | 1 | 45 | 101 |
| Low Soy | 2 | 16 | 72 |
| Low Soy | 2 | 51 | 143 |
| Low Soy | 3 | 3 | 57 |
| Low Soy | 3 | 29 | 81 |
| Low Soy | 3 | 33 | 83 |
| Low Soy | 3 | 46 | 126 |
| Low Soy | 3 | 55 | 155 |
| Low Soy | 4 | 8 | 72 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| High Fish | 146 | 52 | 145 |
| High Fish | 147 | 16 | 78 |
| High Fish | 147 | 33 | 97 |
| High Fish | 146 | 52 | 145 |

In this example, if Time were considered a fixed-effects factor, each time point would be a different level of the factor, yielding too many levels. The appropriate approach in this example is to include Time as a covariate and examine the linear relationship (perhaps following a transformation) between Time and Weight. In this example, the nature of the design requires that Time be a covariate.

Common experiments in which time should be included as a covariate are experiments involving human subjects that don’t report on schedule.

## Using a Time Variable When Time is Not Measured in the Study

Many designs (e.g., factorial, split-plot, ANCOVA) for which the use of mixed models is recommended do not have time as a measured variable. In such cases, it can still be useful to include a time variable as an ordering variable. This is particularly important when the dataset itself is not ordered, when there are missing values, and when the specified covariance structure is complex. An example of a design where time is included only for ordering purposes is a cross-over design.

## A Model-Building Strategy

There are three main components of a mixed model:

- **The Fixed Effects Component.** The fixed effects component of the model consists of the fixed factors, the covariates, and the interactions of fixed factors and covariates. The strength of evidence for the true effect of each fixed effects term is given by the probability level of the corresponding F-test.
- **The Random (Subject) Component.** The random factors include all random factors and (possibly) interactions of random factors with fixed factor variables or covariates. The importance of each random term is more subjective. Inclusion or exclusion of a random term is often decided by comparing the magnitude of the estimates. Relatively small estimates may, in some cases, be removed from the model. The meaning of 'relatively small' is beyond the scope of this manual.
- **The Covariance Pattern of Repeated Measurements.** The covariance pattern indicates the pattern of the residual error of repeated measurements. Specific patterns are shown in detail later in this chapter. The pattern should usually be Diagonal if a random model is specified. Patterns can be compared by examining the AIC value for each pattern. A separate run is required for each pattern.

The underlying goal in building a mixed model should be finding the simplest model that best fits the observed data. A reasonable top-down strategy for building a model might include the following steps:

1. Specify all the fixed effects, covariates, and potentially important interactions in the Fixed Effects Model.
2. Specify either the Random Model or the Repeated Covariance Pattern as the circumstances dictate.
3. Run the model.
4. Compare the random terms to see if any are clearly negligible (e.g., less than 20 times smaller than the others).
5. Re-run the model excluding the negligible random terms.
6. Examine the fixed effects terms F-tests tests. Iteratively remove interaction terms from the fixed effects model that have large probability levels until all are below, say, 0.20.
7. If a Repeated Covariance Pattern is of interest, re-run the analysis several times with different patterns, comparing the AIC values. Keep the pattern with the lowest AIC value.
8. Run the final model with comparisons of interest and specific covariate values.

This strategy is one among many that could be used in refining a mixed model. In some cases, regulations may dictate the terms that may or may not be included in the model, which leaves little or no room for refinement. The order of steps given here is subjective, but perhaps gives a feel for the considerations that should be made in determining a good model. The discussion near the end of Example 1 involving model refinement for a specific example may also be helpful.

# Multiple Comparisons of Fixed Effect Levels

If there is evidence that a fixed factor of a mixed model has difference responses among its levels, it is usually of interest to perform post-hoc pair-wise comparisons of the least-squares means to further clarify those differences. It is well-known that p-value adjustments need to be made when multiple tests are performed (see Hochberg and Tamhane, 1987, or Hsu, 1996, for general discussion and details of the need for multiplicity adjustment). Such adjustments are usually made to preserve the family-wise error rate (FWER), also called the experiment-wise error rate, of the group of tests. FWER is the probability of incorrectly rejecting at least one of the pair-wise tests.

## Family-Wise Error Rate (FWER) Control – Bonferroni Adjustment

The Bonferroni p-value adjustment produces adjusted p-values (probability levels) for which the FWER is controlled strictly (Westfall et al, 1999). The Bonferroni adjustment is applied to all $m$ unadjusted (raw) p-values ( $p_j$ ) as

$$\tilde{p}_j = \min\left(mp_j, 1\right).$$

That is, each p-value is multiplied by the number of tests in the set (family), and if the result is greater than one, it is set to the maximum possible p-value of one.

The Bonferroni adjustment is generally considered to be a conservative method for simultaneously comparing levels of fixed effects.

In the following example, four levels of a fixed factor are compared (all pairs): A, B, C, and D.

**Multiple Comparison Example – Main Effects**

| Test | Raw P-value | Bonferroni Adjusted P-value |
|---|---|---|
| A vs B | 0.01435 | 0.08610 |
| A vs C | 0.00762 | 0.04572 |
| A vs D | 0.00487 | 0.02922 |
| B vs C | 0.34981 | 1.00000 |
| B vs D | 0.06062 | 0.36372 |
| C vs D | 0.71405 | 1.00000 |

In this example, the adjustments are based on m = 6 tests.

## Multiple Comparisons for the Interaction of Two Main Effects

When examining a fixed effect interaction using post-hoc (or planned) multiple comparison tests, a useful method is to compare all levels of one factor at each level of the other factor. This method is termed ‘slicing’. For example, if the interaction of Time and Treatment is significant, comparing the treatment levels at each time point could aid in understanding the nature of the interaction.

**Multiple Comparison Example – Interaction**

| Time | Test | Raw P-value | Bonferroni Adjusted P-value |
|---|---|---|---|
| 1 hour | A vs B | 0.25186 | 1.00000 |
| 1 hour | A vs C | 0.00118 | 0.02124 |
| 1 hour | A vs D | 0.13526 | 1.00000 |
| 1 hour | B vs C | 0.07275 | 1.00000 |
| 1 hour | B vs D | 0.12994 | 1.00000 |
| 1 hour | C vs D | 0.08068 | 1.00000 |
| 5 hours | A vs B | 0.11279 | 1.00000 |
| 5 hours | A vs C | 0.01779 | 0.32022 |
| 5 hours | A vs D | 0.18634 | 1.00000 |
| 5 hours | B vs C | 0.07291 | 1.00000 |
| 5 hours | B vs D | 0.05254 | 0.94572 |
| 5 hours | C vs D | 0.03883 | 0.69894 |
| 10 hours | A vs B | 0.14701 | 1.00000 |
| 10 hours | A vs C | 0.02798 | 0.50364 |
| 10 hours | A vs D | 0.15722 | 1.00000 |
| 10 hours | B vs C | 0.13614 | 1.00000 |
| 10 hours | B vs D | 0.10642 | 1.00000 |
| 10 hours | C vs D | 0.16751 | 1.00000 |

In this example, the adjustments are based on m = 18 tests. It can be seen from this example that minimizing the number of tests enhances the power to detect significant differences.

## Multiple Comparisons for Several Covariate Levels

When more than one covariate value is specified for 'Compute Means at these Values' on the Covariates tab, the number of test used in the Bonferroni adjustment can increase dramatically. The number of tests for the Bonferroni adjustment is computed as

$$\textit{Number of Tests} = \textit{Number of Comparisons per Set} \times \textit{Number of Covariate Sets}$$

As an example, suppose that an experiment has two covariates, and a single fixed treatment factor with three levels: Control, T1, and T2. If 'All Pairs' were selected as the comparison on the Comparisons tab, then the number of comparisons per set would be three (T1 – Control, T2 – Control, and T2 – T1). Suppose that the researcher desired to compute the hypothesis tests at two values for the first covariate and four values for the second. The number of covariate sets would be $2 \times 4 = 8$. Therefore, the number of tests used in the Bonferroni adjustment to conserve the overall error-rate would be $3 \times 8 = 24$. The raw p-value would have to be less than $0.05/24 = 0.00208$ in order to declare significance at the 0.05 level.

This example illustrates that care must be taken when specifying the covariate values at which the means and analyses will be computed. As more covariate values are specified, the number of tests in the adjustment increases making it more and more difficult to find differences that are significant.

# Mixed Model Technical Details

As stated previously, the general form of the linear mixed model is

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}$$

where

- $\mathbf{y}$ vector of responses
- $\mathbf{X}$ known design matrix of the fixed effects
- $\boldsymbol{\beta}$ unknown vector of fixed effects parameters to be estimated
- $\mathbf{Z}$ known design matrix of the random effects
- $\mathbf{u}$ unknown vector of random effects
- $\boldsymbol{\varepsilon}$ unobserved vector of random errors

We assume

$\mathbf{u} \sim N(\mathbf{0},\mathbf{G})$

$\boldsymbol{\varepsilon} \sim N(\mathbf{0},\mathbf{R})$

$\text{Cov}[\mathbf{u}, \boldsymbol{\varepsilon}] = \mathbf{0}$

where

- $\mathbf{G}$ variance-covariance matrix of $\mathbf{u}$
- $\mathbf{R}$ variance-covariance matrix of the errors $\boldsymbol{\varepsilon}$

The variance of $\mathbf{y}$, denoted $\mathbf{V}$, is

$$\begin{aligned}\mathbf{V} &= \text{Var}[\mathbf{y}] \\ &= \text{Var}[\mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}] \\ &= 0 + \text{Var}[\mathbf{Zu} + \boldsymbol{\varepsilon}] \\ &= \mathbf{ZGZ'} + \mathbf{R}\end{aligned}$$

In order to test the parameters in $\boldsymbol{\beta}$, which is typically the goal in mixed model analysis, the unknown parameters ($\boldsymbol{\beta}$, $\mathbf{G}$, and $\mathbf{R}$) must be estimated. Estimates for $\boldsymbol{\beta}$ require estimates of $\mathbf{G}$ and $\mathbf{R}$. In order to estimate $\mathbf{G}$ and $\mathbf{R}$, the structure of $\mathbf{G}$ and $\mathbf{R}$ must be specified. Structures for $\mathbf{G}$ and $\mathbf{R}$ are discussed later.

## Individual Subject Formulation

Because of the size of the matrices that are involved in mixed model analysis, it is useful for computational purposes to reduce the dimensionality of the problem by analyzing the data one subject at a time. Because the data from different subjects are statistically independent, the log-likelihood of the data can be summed over the subjects, according to the formulas below. Before we look at the likelihood functions, we examine the linear mixed model for a particular subject:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, N$$

where

- $\mathbf{y}_i$ $n_i \times 1$ vector of responses for subject $i$.

$\mathbf{X}_i$ $n_i \times p$ design matrix of fixed effects for subject $i$ ($p$ is the number of columns in $\mathbf{X}$).

$\boldsymbol{\beta}$ $p \times 1$ vector of regression parameters.

$\mathbf{Z}_i$ $n_i \times q$ design matrix of the random effects for subject $i$.

$\mathbf{u}_i$ $q \times 1$ vector of random effects for subject $i$ which has means of zero and covariance matrix $\mathbf{G}_{sub}$.

$\boldsymbol{\varepsilon}_i$ $n_i \times 1$ vector of errors for subject $i$ with zero mean and covariance $\mathbf{R}_i$.

$n_i$ number of repeated measurements on subject $i$.

$N$ number of subjects.

The following definitions will also be useful.

$\mathbf{e}_i$ vector of residuals for subject $i$ ($\mathbf{e}_i = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}$).

$\mathbf{V}_i$ $\text{Var}[\mathbf{y}_i] = \mathbf{Z}_i\mathbf{G}_{sub}\mathbf{Z}_i' + \mathbf{R}_i$

To see how the individual subject mixed model formulation relates to the general form, we have

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_N \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_N \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_N \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{pmatrix}$$

## Likelihood Formulas

Rather than maximizing the likelihood function, it is convenient (for theoretical and practical reasons beyond the scope of this manual) to minimize -2 times the log likelihood function rather than maximize the likelihood function itself. There are two types of likelihood estimation methods that are generally considered in mixed model estimation: maximum likelihood (ML) and restricted maximum likelihood (REML). REML is generally favored over ML because the variance estimates using REML are unbiased for small sample sizes, whereas ML estimates are unbiased only asymptotically (see Littell et al., 2006 or Demidenko, 2004). Both estimation methods are available in ***NCSS***.

### Maximum Likelihood

The general form -2 log-likelihood ML function is

$$-2L_{ML}(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}) = \ln|\mathbf{V}| + \mathbf{e}'\mathbf{V}^{-1}\mathbf{e} + N_T \ln(2\pi)$$

The equivalent individual subject form is

$$-2L_{\mathbf{ML}}(\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}) = \sum_{\mathbf{i}=1}^{N}\left(\ln|\mathbf{V_i}| + \mathbf{e_i'}\mathbf{V_i^{-1}}\mathbf{e_i}\right) + \mathbf{N_T}\ln(2\pi)$$

where $N_T$ is the total number of observations, or

$$N_T = \sum_{i=1}^{N} n_i$$

## Restricted Maximum Likelihood

The general form -2 log-likelihood REML function is

$$-2L_{REML}(\boldsymbol{\beta},\mathbf{G},\mathbf{R}) = \ln|\mathbf{V}| + \mathbf{e}'\mathbf{V}^{-1}\mathbf{e} + \ln\left|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right| + (N_T - p)\ln(2\pi)$$

The equivalent individual subject form is

$$-2L_{REML}(\boldsymbol{\beta},\mathbf{G},\mathbf{R}) = \sum_{i=1}^{N}\left[\ln|\mathbf{V}_i| + \mathbf{e}_i'\mathbf{V}_i^{-1}\mathbf{e}_i\right] + \ln\left|\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\right| + (N_T - p)\ln(2\pi)$$

where, again, $N_T$ is the total number of observations, or

$$N_T = \sum_{i=1}^{N} n_i$$

and $p$ is the number of columns in $\mathbf{X}$ or $\mathbf{X}_i$.

## The G Matrix

The **G** matrix is the variance-covariance matrix for the random effects **u**. Typically, when the **G** matrix is used to specify the variance-covariance structure of **y**, the structure for **R** is simply $\sigma^2\mathbf{I}$. Caution should be used when both **G** and **R** are specified as complex structures, since large numbers of sometimes redundant covariance elements can result.

The **G** matrix is made up of $N$ symmetric $\mathbf{G}_{sub}$ matrices,

$$\mathbf{G} = \begin{pmatrix} G_{sub} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & G_{sub} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & G_{sub} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & G_{sub} \end{pmatrix}$$

The dimension of $\mathbf{G}_{sub}$ is $q \times q$, where $q$ is the number of random effects for each subject.

### Structures of $\mathbf{G}_{sub}$

There are two commonly used structures for the elements of the $\mathbf{G}_{sub}$ matrix: diagonal and unstructured.

**Diagonal $\mathbf{G}_{sub}$**

$$\mathbf{G}_{sub} = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{pmatrix}$$

**Unstructured $\mathbf{G}_{sub}$**

$$\mathbf{G}_{sub} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

The diagonal $\mathbf{G}_{sub}$ should be used when there is no covariance between parameters, such as in the random effects models. The unstructured $\mathbf{G}_{sub}$ is typically used when you want to include covariances, such as in random coefficients models.

## The R Matrix

The **R** matrix is the variance-covariance matrix for errors, **ε**. When the **R** matrix is used to specify the variance-covariance structure of **y**, the $\mathbf{G}_{sub}$ matrix is not used.

The full **R** matrix is made up of $N$ symmetric **R** sub-matrices,

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_N \end{pmatrix}$$

where $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \cdots, \mathbf{R}_N$ are all of the same structure, but, unlike the $\mathbf{G}_{sub}$ matrices, differ according to the number of repeated measurements on each subject.

When the **R** matrix is specified in ***NCSS***, it is assumed that there is a fixed, known set of repeated measurement times. (If the repeated measurement times are random, specification of the $\mathbf{G}_{sub}$ matrix with $\mathbf{R} = \sigma^2\mathbf{I}$ should be used instead for specifying covariance structure.) Thus, the differences in the dimensions of the **R** sub-matrices occur only when some measurements for a subject are missing.

As an example, suppose an **R** sub-matrix is of the form

$$\mathbf{R}_{Sub} = \begin{pmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \sigma_3^2 & & \\ & & & \sigma_4^2 & \\ & & & & \sigma_5^2 \end{pmatrix},$$

where there are five time points at which each subject is intended to be measured: 1 hour, 2 hours, 5 hours, 10 hours, and 24 hours. If the first subject has measurements at all five time points, then $n_1 = 5$, and the sub-matrix is identical to $\mathbf{R}_{Sub}$ above, and $\mathbf{R}_1 = \mathbf{R}_{Sub}$.

Suppose the second subject is measured at 1 hour, 5 hours, and 24 hours, but misses the 2-hour and 10-hour measurements. The $\mathbf{R}_2$ matrix for this subject is

$$\mathbf{R}_2 = \begin{pmatrix} \sigma_1^2 & & \\ & \sigma_3^2 & \\ & & \sigma_5^2 \end{pmatrix}.$$

For this subject, $n_2 = 3$. That is, for the case when the time points are fixed, instead of having missing values in the **R** sub-matrices, the matrix is collapsed to accommodate the number of realized measurements.

## Structures of R

There are many possible structures for the sub-matrices that make up the **R** matrix. The $\mathbf{R}_{Sub}$ structures that can be specified in ***NCSS*** are shown below.

### Diagonal

Homogeneous

$$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

### Compound Symmetry

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \rho\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho\sigma_4\sigma_1 & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$$

### AR(1)

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 \\ \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \rho^3\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho^2\sigma_2\sigma_4 \\ \rho^2\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho^3\sigma_4\sigma_1 & \rho^2\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

**Toeplitz**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 & \rho_3\sigma^2 \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ \rho_3\sigma^2 & \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \rho_3\sigma_1\sigma_4 \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 \\ \rho_2\sigma_3\sigma_1 & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ \rho_3\sigma_4\sigma_1 & \rho_2\sigma_4\sigma_2 & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{pmatrix}$$

**Toeplitz(2)**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & & \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \\ & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ & & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & & \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \\ & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ & & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho_1 & & \\ \rho_1 & 1 & \rho_1 & \\ & \rho_1 & 1 & \rho_1 \\ & & \rho_1 & 1 \end{pmatrix}$$

Note: This is the same as Banded(2).

**Toeplitz(3)**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 & \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ & \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 \\ \rho_2\sigma_3\sigma_1 & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ & \rho_2\sigma_4\sigma_2 & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ & \rho_2 & \rho_1 & 1 \end{pmatrix}$$

**Toeplitz(4) and Toeplitz(5)**

Toeplitz(4) and Toeplitz(5) follow the same pattern as Toeplitz(2) and Toeplitz(3), but with the corresponding numbers of bands.

**Banded(2)**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & & \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \\ & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ & & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & & \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \\ & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ & & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho & & \\ \rho & 1 & \rho & \\ & \rho & 1 & \rho \\ & & \rho & 1 \end{pmatrix}$$

Note: This is the same as Toeplitz(1).

**Banded(3)**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho & \rho & \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ & \rho & \rho & 1 \end{pmatrix}$$

**Banded(4) and Banded (5)**

Banded(4) and Banded(5) follow the same pattern as Banded(2) and Banded(3), but with the corresponding numbers of bands.

**Unstructured**

Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_{12}\sigma^2 & \rho_{13}\sigma^2 & \rho_{14}\sigma^2 \\ \rho_{21}\sigma^2 & \sigma^2 & \rho_{23}\sigma^2 & \rho_{24}\sigma^2 \\ \rho_{31}\sigma^2 & \rho_{32}\sigma^2 & \sigma^2 & \rho_{34}\sigma^2 \\ \rho_{41}\sigma^2 & \rho_{42}\sigma^2 & \rho_{43}\sigma^2 & \sigma^2 \end{pmatrix}$$

Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 \\ \rho_{31}\sigma_3\sigma_1 & \rho_{32}\sigma_3\sigma_2 & \sigma_3^2 & \rho_{34}\sigma_3\sigma_4 \\ \rho_{41}\sigma_4\sigma_1 & \rho_{42}\sigma_4\sigma_2 & \rho_{43}\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

Correlation

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}$$

## Partitioning the Variance-Covariance Structure with Groups

In the case where it is expected that the variance-covariance parameters are different across group levels of the data, it may be useful to specify a different set of **R** or **G** parameters for each level of a group variable. This produces a set of variance-covariance parameters that is different for each level of the chosen group variable, but each set has the same structure as the other groups.

## Partitioning the G Matrix Parameters

Suppose the structure of **G** is specified to be diagonal. If $\mathbf{G}_{sub}$ has four parameters then

$$\mathbf{G}_{sub} = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{pmatrix}.$$

If there are twenty subjects, then

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{sub} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{sub} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{sub} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_{sub} \end{pmatrix}.$$

The total number of variance parameters is four.

Suppose now that there are two groups of ten subjects, and it is believed that the four variance parameters of the first group are different from the four variance parameters of the second group.

We now have

$$\mathbf{G}_1 = \begin{pmatrix} \sigma_{11}^2 & & & \\ & \sigma_{12}^2 & & \\ & & \sigma_{13}^2 & \\ & & & \sigma_{14}^2 \end{pmatrix}, \text{ and } \mathbf{G}_2 = \begin{pmatrix} \sigma_{21}^2 & & & \\ & \sigma_{22}^2 & & \\ & & \sigma_{23}^2 & \\ & & & \sigma_{24}^2 \end{pmatrix}.$$

If the first ten subjects are in Group 1, then the **G** matrix becomes

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_1 & & & & & & & \\ & \mathbf{G}_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \mathbf{G}_1 & & & & \\ & & & & \mathbf{G}_2 & & & \\ & & & & & \mathbf{G}_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \mathbf{G}_2 \end{pmatrix},$$

with eight variance parameters, rather than four.

## Partitioning the R Matrix Parameters

Suppose the structure of **R** in a study with four time points is specified to be Toeplitz:

$$\mathbf{R} = \begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 & \rho_3\sigma^2 \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ \rho_3\sigma^2 & \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}.$$

If there are sixteen subjects then

$$\mathbf{R} = \begin{pmatrix} \mathbf{R_1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R_2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_3} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R_{16}} \end{pmatrix}.$$

The total number of variance-covariance parameters is four: $\sigma^2, \rho_1, \rho_2,$ and $\rho_3$.

Suppose now that there are two groups of eight subjects, and it is believed that the four variance parameters of the first group are different from the four variance parameters of the second group.

We now have

$$\mathbf{R}_1,\cdots,\mathbf{R}_8 = \begin{pmatrix} \sigma_1^2 & \rho_{11}\sigma^2 & \rho_{12}\sigma^2 & \rho_{13}\sigma^2 \\ \rho_{11}\sigma^2 & \sigma_1^2 & \rho_{11}\sigma^2 & \rho_{12}\sigma^2 \\ \rho_{12}\sigma^2 & \rho_{11}\sigma^2 & \sigma_1^2 & \rho_{11}\sigma^2 \\ \rho_{13}\sigma^2 & \rho_{12}\sigma^2 & \rho_{11}\sigma^2 & \sigma_1^2 \end{pmatrix},$$

and

$$\mathbf{R}_9,\cdots,\mathbf{R}_{16} = \begin{pmatrix} \sigma_2^2 & \rho_{21}\sigma^2 & \rho_{22}\sigma^2 & \rho_{23}\sigma^2 \\ \rho_{21}\sigma^2 & \sigma_2^2 & \rho_{21}\sigma^2 & \rho_{22}\sigma^2 \\ \rho_{22}\sigma^2 & \rho_{21}\sigma^2 & \sigma_2^2 & \rho_{21}\sigma^2 \\ \rho_{23}\sigma^2 & \rho_{22}\sigma^2 & \rho_{21}\sigma^2 & \sigma_2^2 \end{pmatrix}.$$

The total number of variance-covariance parameters is now eight.

It is easy to see how quickly the number of variance-covariance parameters increases when **R** or **G** is partitioned by groups.

### Repeated Measures Complication in Partitioning R

When partitioning the variance-covariance parameters into groups in some less-common repeated-measures designs, more than one group can occur within a subject. Re-examining the **R** partitioning example above, suppose instead that all sixteen subjects are measured four times: twice with Treatment A, and twice with Treatment B. For the sake of this example, assume that the first eight subjects receive A, A, B, B and the second eight receive B, B, A, A. The covariance parameters across treatments but within a subject are assumed to be zero, and the **R** sub-matrices for the first eight subjects become

$$\mathrm{R}_1,\cdots,\mathrm{R}_8 = \begin{pmatrix} \sigma_\mathrm{A}^2 & \rho_\mathrm{A}\sigma_\mathrm{A}^2 & & \\ \rho_\mathrm{A}\sigma_\mathrm{A}^2 & \sigma_\mathrm{A}^2 & & \\ & & \sigma_\mathrm{B}^2 & \rho_\mathrm{B}\sigma_\mathrm{B}^2 \\ & & \rho_\mathrm{B}\sigma_\mathrm{B}^2 & \sigma_\mathrm{B}^2 \end{pmatrix},$$

and for the last eight subjects,

$$\mathrm{R}_9,\cdots,\mathrm{R}_{16} = \begin{pmatrix} \sigma_\mathrm{B}^2 & \rho_\mathrm{B}\sigma_\mathrm{B}^2 & & \\ \rho_\mathrm{B}\sigma_\mathrm{B}^2 & \sigma_\mathrm{B}^2 & & \\ & & \sigma_\mathrm{A}^2 & \rho_\mathrm{A}\sigma_\mathrm{A}^2 \\ & & \rho_\mathrm{A}\sigma_\mathrm{A}^2 & \sigma_\mathrm{A}^2 \end{pmatrix}.$$

The total number of variance-covariance parameters is only four: $\sigma_\mathrm{A}^2, \sigma_\mathrm{B}^2, \rho_\mathrm{A},$ and $\rho_\mathrm{B}$.

In general, when we attempt to divide the variance-covariance parameters into groups with a repeated-measures design, the covariance of residuals within a subject, but across treatments, is assumed to be zero.

# Estimating and Testing Fixed Effects Parameters

The estimation phase in the analysis of a mixed model produces variance and covariance parameter estimates of the elements of **G** and **R**, giving $\hat{\mathbf{R}}$ and $\hat{\mathbf{G}}$, and hence, $\hat{\mathbf{V}}$. The REML and ML solutions for $\hat{\boldsymbol{\beta}}$ are given by

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{y}$$

with estimated variance-covariance

$$\hat{\Sigma} = \text{var}\left(\hat{\boldsymbol{\beta}}\right) = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}$$

See, for example, Brown and Prescott (2006), Muller and Stewart (2006), or Demidenko (2004) for more details of the estimating equations.

Hypothesis tests and confidence intervals for $\boldsymbol{\beta}$ are formed using a linear combination matrix (or vector) **L**.

## L Matrix Details

**L** matrices specify linear combinations of $\boldsymbol{\beta}$ corresponding to means or hypothesis tests of interest. Essentially, the **L** matrix defines the mean or test. The number of columns in each **L** matrix is the same as the number of elements of $\boldsymbol{\beta}$. For estimating a particular mean, the **L** matrix consists of a single row. For hypothesis tests, the number of rows of **L** varies according to the test. Below are some examples of **L** matrices that arise in common analyses:

### L Matrix for Testing a Single Factor (Food with 4 levels) in a Single-Factor Model

| No. | Effect | Food | L1 | L2 | L3 |
|---|---|---|---|---|---|
| 1 | Intercept | | | | |
| 2 | Food | HighIron | 1.0000 | 1.0000 | 1.0000 |
| 3 | Food | LowIron | -1.0000 | | |
| 4 | Food | None | | -1.0000 | |
| 5 | Food | Salicyl | | | -1.0000 |

### L Matrix for a Single Mean (LowIron) of a Single Factor (4 levels) in a Single-Factor Model

| No. | Effect | Food | L1 |
|---|---|---|---|
| 1 | Intercept | | 1.0000 |
| 2 | Food | HighIron | |
| 3 | Food | LowIron | 1.0000 |
| 4 | Food | None | |
| 5 | Food | Salicyl | |

## L Matrix for Testing a Single Factor (Drug – 3 levels) in a Two-Factor Model with Interaction

| No. | Effect | Drug | Time | L1 | L2 |
|---|---|---|---|---|---|
| 1 | Intercept | | | | |
| 2 | Drug | Kerlosin | | 1.0000 | 1.0000 |
| 3 | Drug | Laposec | | -1.0000 | |
| 4 | Drug | Placebo | | | -1.0000 |
| 5 | Time | | 0.5 | | |
| 6 | Time | | 1 | | |
| 7 | Time | | 1.5 | | |
| 8 | Time | | 2 | | |
| 9 | Time | | 2.5 | | |
| 10 | Time | | 3 | | |
| 11 | Drug*Time | Kerlosin | 0.5 | 0.1667 | 0.1667 |
| 12 | Drug*Time | Kerlosin | 1 | 0.1667 | 0.1667 |
| 13 | Drug*Time | Kerlosin | 1.5 | 0.1667 | 0.1667 |
| 14 | Drug*Time | Kerlosin | 2 | 0.1667 | 0.1667 |
| 15 | Drug*Time | Kerlosin | 2.5 | 0.1667 | 0.1667 |
| 16 | Drug*Time | Kerlosin | 3 | 0.1667 | 0.1667 |
| 17 | Drug*Time | Laposec | 0.5 | -0.1667 | |
| 18 | Drug*Time | Laposec | 1 | -0.1667 | |
| 19 | Drug*Time | Laposec | 1.5 | -0.1667 | |
| 20 | Drug*Time | Laposec | 2 | -0.1667 | |
| 21 | Drug*Time | Laposec | 2.5 | -0.1667 | |
| 22 | Drug*Time | Laposec | 3 | -0.1667 | |
| 23 | Drug*Time | Placebo | 0.5 | | -0.1667 |
| 24 | Drug*Time | Placebo | 1 | | -0.1667 |
| 25 | Drug*Time | Placebo | 1.5 | | -0.1667 |
| 26 | Drug*Time | Placebo | 2 | | -0.1667 |
| 27 | Drug*Time | Placebo | 2.5 | | -0.1667 |
| 28 | Drug*Time | Placebo | 3 | | -0.1667 |

## L Matrix for Testing a Covariate in a One-Factor (3 levels) Model with a Covariate

| No. | Effect | Drug | L1 |
|---|---|---|---|
| 1 | Intercept | | |
| 2 | Drug | Kerlosin | |
| 3 | Drug | Laposec | |
| 4 | Drug | Placebo | |
| 5 | Weight | | 1.0000 |
| 6 | Drug*Weight | Kerlosin | 0.3333 |
| 7 | Drug*Weight | Laposec | 0.3333 |
| 8 | Drug*Weight | Placebo | 0.3333 |

## Kenward and Roger Fixed Effects Hypothesis Tests

Hypothesis tests have the general form

$$H_0: \mathbf{L}\boldsymbol{\beta} = \mathbf{0}$$

where $\mathbf{L}$ is a linear contrast matrix of rank $h$ corresponding to the desired comparisons to be made in the hypothesis test. Let $d$ be the denominator degrees of freedom and $q$ be the number of variance-covariance parameters, which is the dimension of $\mathbf{W}$ (defined below).

The Kenward and Roger (1997) test statistic for testing $H_0$ is

$$\mathbf{F}_{\mathbf{h,d}} = \frac{\lambda}{h}\hat{\boldsymbol{\beta}}'\mathbf{L}'\left(\mathbf{L}\mathbf{C}^{*}\mathbf{L}'\right)^{-1}\mathbf{L}\hat{\boldsymbol{\beta}}$$

where

$$\mathbf{C}^{*} = \mathbf{C} + 2\mathbf{C}\left\{\sum_{\mathbf{r}=1}^{\mathbf{q}}\sum_{\mathbf{s}=1}^{\mathbf{q}}\mathbf{W}_{\mathbf{rs}}\left(\mathbf{Q}_{\mathbf{rs}} - \mathbf{P}_{\mathbf{r}}\mathbf{C}\mathbf{P}_{\mathbf{s}} - \frac{1}{4}\mathbf{S}_{\mathbf{rs}}\right)\right\}\mathbf{C}$$

$$\mathbf{C} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}$$

$$\mathbf{Q}_{\mathbf{rs}} = \mathbf{X}'\mathbf{V}^{-1}\dot{\mathbf{V}}_{\mathbf{r}}\mathbf{V}^{-1}\dot{\mathbf{V}}_{\mathbf{s}}\mathbf{V}^{-1}\mathbf{X} = \sum_{i=1}^{N}\mathbf{X}_{\mathbf{i}}'\mathbf{V}_{\mathbf{i}}^{-1}\dot{\mathbf{V}}_{\mathbf{ri}}\mathbf{V}_{\mathbf{i}}^{-1}\dot{\mathbf{V}}_{\mathbf{si}}\mathbf{V}_{\mathbf{i}}^{-1}\mathbf{X}_{\mathbf{i}}$$

$$\mathbf{P}_{\mathbf{r}} = -\mathbf{X}'\mathbf{V}^{-1}\dot{\mathbf{V}}_{\mathbf{r}}\mathbf{V}^{-1}\mathbf{X} = -\sum_{i=1}^{N}\mathbf{X}_{\mathbf{i}}'\mathbf{V}_{\mathbf{i}}^{-1}\dot{\mathbf{V}}_{\mathbf{ri}}\mathbf{V}_{\mathbf{i}}^{-1}\mathbf{X}_{\mathbf{i}}$$

$$\mathbf{S}_{\mathbf{rs}} = \mathbf{X}'\mathbf{V}^{-1}\ddot{\mathbf{V}}_{\mathbf{rs}}\mathbf{V}^{-1}\mathbf{X} = \sum_{i=1}^{N}\mathbf{X}_{\mathbf{i}}'\mathbf{V}_{\mathbf{i}}^{-1}\ddot{\mathbf{V}}_{\mathbf{rsi}}\mathbf{V}_{\mathbf{i}}^{-1}\mathbf{X}_{\mathbf{i}}$$

$$\mathbf{W} = \mathbf{H}^{-1}$$

$$\{\mathbf{H}\}_{\mathbf{rs}} = \{\text{Hessian}\}_{\mathbf{rs}}$$

$$\dot{\mathbf{V}}_{\mathbf{r}} = \frac{\partial \mathbf{V}}{\partial \sigma_{\mathbf{r}}}$$

$$\ddot{\mathbf{V}}_{\mathbf{rs}} = \frac{\partial^2 \mathbf{V}}{\partial \sigma_{\mathbf{r}}\partial \sigma_{\mathbf{s}}}$$

$$\mathbf{T} = \mathbf{L}'\left(\mathbf{L}\mathbf{C}\mathbf{L}'\right)^{-1}\mathbf{L}$$

$$a_1 = \sum_{\mathbf{r}=1}^{\mathbf{q}}\sum_{\mathbf{s}=1}^{\mathbf{q}}\mathbf{W}_{\mathbf{rs}}\,\mathrm{tr}(\mathbf{T}\mathbf{C}\mathbf{P}_{\mathbf{r}}\mathbf{C})\,\mathrm{tr}(\mathbf{T}\mathbf{C}\mathbf{P}_{\mathbf{s}}\mathbf{C}),\quad a_2 = \sum_{\mathbf{r}=1}^{\mathbf{q}}\sum_{\mathbf{s}=1}^{\mathbf{q}}\mathbf{W}_{\mathbf{rs}}\,\mathrm{tr}(\mathbf{T}\mathbf{C}\mathbf{P}_{\mathbf{r}}\mathbf{C}\mathbf{T}\mathbf{C}\mathbf{P}_{\mathbf{s}}\mathbf{C})$$

$$a_3 = \frac{a_1 + 6a_2}{2h},\ e = \left(1 - \frac{a_2}{h}\right)^{-1},\ v = \frac{2}{h}\left\{\frac{1 + c_1 a_3}{(1 - c_2 a_3)^2(1 - c_3 a_3)}\right\}$$

$$c_1 = \frac{g}{3h + 2(1-g)},\ c_2 = \frac{h - g}{3h + 2(1-g)},\ c_3 = \frac{h + 2 - g}{3h + 2(1-g)},\ c_4 = \frac{v}{2e^2}$$

$$g = \frac{(h+1)a_1 - (h+4)a_2}{(h+2)a_2}$$

$$d = 4 + \frac{h+2}{c_4 h - 1},\ \lambda = \frac{d}{e(d\text{-}2)}$$

## Kenward and Roger Fixed Effects Confidence Intervals

Confidence intervals for linear combinations of $\boldsymbol{\beta}$ are formed as

$$\mathbf{L}\hat{\boldsymbol{\beta}} \pm \mathbf{t}_{\mathbf{m},\alpha/2}\sqrt{\mathbf{LCL}'}$$

where $t_{m,\alpha/2}$ is the 1-$\alpha$/2 percentile of the $t$ distribution with $m$ degrees of freedom, with

$\mathbf{C}$ and $m$ defined above.

# Solution Algorithms

## Methods for Finding Likelihood Solutions (Newton-Raphson, Fisher Scoring, MIVQUE, and Differential Evolution)

There are four techniques in the Mixed Models procedure for determining the maximum likelihood or restricted maximum likelihood solution (optimum): Newton-Raphson, Fisher Scoring, MIVQUE, and Differential Evolution.

The general steps for the Newton-Raphson, Fisher Scoring, and Differential Evolution techniques are (let $\boldsymbol{\theta}$ be the overall covariance parameter vector):

1. Roughly estimate $\boldsymbol{\theta}$ according to the specified structure for each.
2. Evaluate the likelihood of the model given the data and the estimates of $\boldsymbol{\theta}$.
3. Improve upon the estimates of $\boldsymbol{\theta}$ using a search algorithm. (Improvement is defined as an increase in likelihood.)
4. Iterate until maximum likelihood is reached, according to some convergence criterion.
5. Use the final $\boldsymbol{\theta}$ estimates to estimate $\boldsymbol{\beta}$.

## Newton-Raphson and Fisher Scoring

The differences in the techniques revolve around the initial estimates in Step 1, and the improvements in estimates made in Step 3. For the Newton-Raphson and Fisher Scoring techniques, Step 3 occurs as follows:

3a. With the estimated $\boldsymbol{\theta}$, compute the gradient vector $\mathbf{g}$, and the Hessian matrix $\mathbf{H}$.
3b. Compute $\mathbf{d} = -\mathbf{H}^{-1}\mathbf{g}$.
3c. Let $\lambda = 1$.
3d. Compute new estimates for $\boldsymbol{\theta}$, iteratively, using $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{i-1} + \lambda\mathbf{d}$.
3e. If $\boldsymbol{\theta}_i$ is a valid set of covariance parameters and improves the likelihood, continue to 3f. Otherwise, reduce $\lambda$ by half and return to Step 3d.
3f. Check for convergence. If the convergence criteria (small change in -2log-likelihood) are met, stop. If the convergence criteria are not met, go back to Step 3a.

The gradient vector $\mathbf{g}$, and the Hessian matrix $\mathbf{H}$, used for the Newton-Raphson and Fisher Scoring techniques for solving the REML equations are shown in the following table:

**REML Gradient (g) and Hessian (H)**

| Technique | Gradient (g) | Hessian (H) |
|---|---|---|
| Newton-Raphson | $\mathbf{g}_1 + \mathbf{g}_2 + \mathbf{g}_3$ | $\mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3$ |
| Fisher Scoring | $\mathbf{g}_1 + \mathbf{g}_2 + \mathbf{g}_3$ | $-\mathbf{H}_1 + \mathbf{H}_3$ |

The gradient vector $\mathbf{g}$, and the Hessian matrix $\mathbf{H}$, used for the Newton-Raphson and Fisher Scoring techniques for solving the ML equations are shown in the following table:

**ML Gradient (g) and Hessian (H)**

| Technique | Gradient (g) | Hessian (H) |
|---|---|---|
| Newton-Raphson | $\mathbf{g}_1 + \mathbf{g}_2$ | $\mathbf{H}_1 + \mathbf{H}_2$ |
| Fisher Scoring | $\mathbf{g}_1 + \mathbf{g}_2$ | $-\mathbf{H}_1$ |

where $\mathbf{g}_1$, $\mathbf{g}_2$, $\mathbf{g}_3$, $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are defined as in Wolfinger, Tobias, and Sall (1994).

### Definitons

$$\dot{\mathbf{V}}_{\mathbf{ri}} = \frac{\partial \mathbf{V_i}}{\partial \boldsymbol{\sigma}_\mathbf{r}}, \quad \ddot{\mathbf{V}}_{\mathbf{rsi}} = \frac{\partial^2 \mathbf{V_i}}{\partial \boldsymbol{\sigma}_\mathbf{r} \partial \boldsymbol{\sigma}_\mathbf{s}}, \mathbf{e_i} = \mathbf{y_i} - \mathbf{X_i}\boldsymbol{\beta}, \quad \mathbf{A_i} = \mathbf{X_i'V_i^{-1}X_i}, \; \mathbf{A} = \sum_{i=1}^{N} \mathbf{X_i'V_i^{-1}X_i} = \sum_{i=1}^{N} \mathbf{A_i},$$

$$\mathbf{C} = \mathbf{A}^{-1}, \quad \dot{\mathbf{A}}_\mathbf{r} = \sum \mathbf{X_i'}\left(\frac{\partial \mathbf{V_i^{-1}}}{\partial \boldsymbol{\sigma}_\mathbf{r}}\right)\mathbf{X_i} = -\sum \mathbf{X_i'}\left(\mathbf{V_i^{-1}}\dot{\mathbf{V}}_{\mathbf{ri}}\mathbf{V_i^{-1}}\right)\mathbf{X_i} = -\mathbf{P_r}$$

$$\mathbf{X}^* = \mathbf{XK}, \quad \mathbf{KK'} = \left(\mathbf{X'V^{-1}X}\right)^{-1}$$

### Likelihoods

$$\mathbf{l}_1 = \frac{1}{2}\sum_{i=1}^{N}\ln\left|\mathbf{V}_i\right|,\quad \mathbf{l}_2 = \frac{1}{2}\sum_{i=1}^{N}\mathbf{e}_i'\mathbf{V}_i^{-1}\mathbf{e}_i,\quad \mathbf{l}_3 = \frac{1}{2}\ln\left|\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i\right| = \frac{1}{2}\ln\left|\sum_{i=1}^{N}\mathbf{A}_i\right| = \frac{1}{2}\ln\left|\mathbf{A}\right|$$

### First Derivatives

$$\mathbf{g}_{1r} = \frac{\partial \mathbf{l}_1}{\partial \boldsymbol{\sigma}_r} = \frac{1}{2}\sum_{i=1}^{N}\mathrm{tr}\left(\mathbf{V}_i^{-1}\dot{\mathbf{V}}_{ri}\right)$$

$$\mathbf{g}_{2r} = \frac{\partial \mathbf{l}_2}{\partial \boldsymbol{\sigma}_r} = -\frac{1}{2}\sum_{i=1}^{N}\mathbf{e}_i'\mathbf{V}_i^{-1}\dot{\mathbf{V}}_{ri}\mathbf{V}_i^{-1}\mathbf{e}_i$$

$$\mathbf{g}_{3r} = \frac{\partial \mathbf{l}_3}{\partial \boldsymbol{\sigma}_r} = -\frac{1}{2}\mathrm{tr}\left[\mathbf{H}_3^r\right]$$

### Second Derivatives

$$\mathbf{H}_{1rs} = \frac{\partial^2 \mathbf{l}_1}{\partial \boldsymbol{\sigma}_r \partial \boldsymbol{\sigma}_s} = -\frac{1}{2}\sum_{i=1}^{N}\left\{\mathrm{tr}\left(\mathbf{V}_i^{-1}\ddot{\mathbf{V}}_{rsi}\right) - \mathrm{tr}\left(\mathbf{V}_i^{-1}\dot{\mathbf{V}}_{ri}\mathbf{V}_i^{-1}\dot{\mathbf{V}}_{si}\right)\right\}$$

$$\mathbf{H}_{2rs} = \frac{\partial^2 \mathbf{l}_2}{\partial \boldsymbol{\sigma}_r \partial \boldsymbol{\sigma}_s} = \frac{1}{2}\left(\mathbf{H}_2^{rs} - 2\mathbf{H}_2^{r'}\mathbf{H}_2^{s}\right)$$

$$\mathbf{H}_{3rs} = \frac{\partial^2 \mathbf{l}_3}{\partial \boldsymbol{\sigma}_r \partial \boldsymbol{\sigma}_s} = \frac{1}{2}\mathrm{tr}\left(\mathbf{H}_3^{rs} - \mathbf{H}_3^{r}\mathbf{H}_3^{s}\right)$$

See Wolfinger, Tobias, and Sall (1994), page 1299, for details.

## MIVQUE

The MIVQUE estimates of $\boldsymbol{\theta}$ in REML estimation are found by solving

$$-(\mathbf{H}_1 + \mathbf{H}_3)\boldsymbol{\theta} = -\mathbf{g}_2.$$

The MIVQUE estimates of $\boldsymbol{\theta}$ in ML estimation are found by solving

$$-\mathbf{H}_1\boldsymbol{\theta} = -\mathbf{g}_2.$$

See Wolfinger, Tobias, and Sall (1994), page 1306, for details.

## Differential Evolution

The differential evolution techniques used in the Mixed Models procedure for the ML and REML optimization are described in Price, Storn, and Lampinen (2005).

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

These panels specify the variables used in the analysis, the solution type, and the model.

### Response Variable

#### Response Variable

This variable contains the numeric responses (measurements) for each of the subjects. There is one measurement per subject per time point. Hence, all responses are in a single column (variable) of the spreadsheet.

### Subject Variable

#### Subject Variable

This variable contains an identification value for each subject. Each subject must have a unique identification number (or name). In a repeated measures design, several measurements are made on each subject.

### Time Variable

#### Time Variable

This variable contains the time at which each measurement is made. If this variable is omitted, the time values are assigned sequentially with the first value being '1', the next value being '2', and so on.

### Factor Variables

#### Factor (Categorical) Variables

Designate any factor (categorical or class) independent variables here. These variables can then be used in the model portion of the Fixed and Random specifications. Note that placing a variable here does NOT automatically include it in a model.

By categorical we mean that the variable has only a few unique values (text or numeric) which are used to identify the categories. Capitalization is ignored when determining unique text values.

### Covariate Variables

#### Covariate (Continuous) Variables

Designate any numeric (continuous) independent variables here. When these variables are included in the Fixed Model statement, the technique is known as Analysis of Covariance (or ANCOVA).

'Numeric' means that the values are at least ordinal. Nominal variables should be specified as Categorical, even though their values may be numeric.

When Covariates are specified, the options on the Covariates tab should be specified for them.

## Options

### Likelihood Type

Specify the type of likelihood equation to be solved. The options are:

- **MLE**

  The 'Maximum Likelihood' solution has become less popular.

- **REML (recommended)**

  The 'Restricted Maximum Likelihood' solution is recommended. It is the default in other software programs (such as SAS).

### Solution Method

Specify the method to be used to solve the likelihood equations. The options are:

- **Newton-Raphson**

  This is an implementation of the popular 'gradient search' procedure for maximizing the likelihood equations. Whenever possible, we recommend that you use this method.

- **Fisher-Scoring**

  This is an intermediate step in the Newton-Raphson procedure. However, when the Newton-Raphson fails to converge, you may want to stop with this procedure.

- **MIVQUE**

  This non-iterative method is used to provide starting values for the Newton-Raphson method. For large problems, you may want to investigate the model using this method since it is much faster.

- **Differential Evolution**

  This grid search technique will often find a solution when the other methods fail to converge. However, it is painfully slow--often requiring hours to converge--and so should only be used as a last resort.

- **Read in from a Variable**

  Use this option when you want to use a solution from a previous run or from another source. The solution is read in from the variable selected in the 'Read Solution From' variable.

### Read Solution From (Variable)

This optional variable contains the variance-covariance parameter values of a solution that has been found previously. The order of the parameter values is the same as on the parameter reports.

This option is useful when problem requires a great deal of time to solve. Once you have achieved a solution, you can reuse it by entering this variable here and setting the 'Solution Method' option to 'Read in from a Variable'.

### Write Solution To (Variable)

Select an empty variable into which the solution is automatically stored. Note that any previous information in this variable will be destroyed.

This option is useful when problem requires a great deal of time to solve. Once you have achieved a solution, you can then reuse it by entering this variable in the 'Read Solution From' variable box and setting the 'Solution Method' option to 'Read in from a Variable'.

### Force Covariance to be Positive

When checked, this option forces all covariances (and correlations) in the Random Components (off-diagonal elements of the G matrix) and Repeated Components (off-diagonal elements of the R matrix) to be non-negative. When this option is not checked, some covariances can be negative.

It usually makes good sense to force these covariances (and thus the corresponding correlations) to be positive. However, occasionally you may want to allow negative covariances.

## Fixed Effects Model

### Model

Specify the statistical model for fixed effects here. Statistical hypothesis tests will be generated for each term in this model. Variables for which hypothesis tests are to be performed should be included in this model statement. You may also include variables in this model that are solely to be used for adjustment and not important for inference or hypothesis testing. For categorical factors, each term represents a set of indicator variables in the expanded design matrix.

The components of this model come from the variables listed in the Factor and Covariate variables. If you want to use them, they must be listed there.

#### Syntax

In the examples that follow each syntax description, 'A', 'B', 'C', and 'D' represent variable names. We will assume that A, B, and C are categorical variables, and D is a covariate.

1. Specify main effects by specifying their variable names on the database, separated by blanks or the '+' (plus) sign.

   | | |
   |---|---|
   | A+B | Main effects for A and B only |
   | A B C | Main effects for A, B, and C only |
   | A B D | Main effects for A and B, plus the covariate effect of D |

2. Specify interactions and cross products using an asterisk (*) between variable names, such as Fruit*Nuts or A*B*C. When an interaction between a discrete factor and a covariate is specified, a cross-product is generated for each value of the factor. For covariates, higher order (e.g. squared, cubic) terms may be added by repeating the covariate name. If D is a covariate, D*D represents the covariate squared, and D*D*D represents the covariate cubed, etc. Only covariates should be repeated. Note that categorical terms should not be squared or cubed. That is, if A is a categorical variable, you would not include A*A nor A*A*A in your model.

   | | |
   |---|---|
   | A+B+A*B | Main effects for A and B plus the AB interaction |
   | A+B+C+A*B+A*C+B*C+A*B*C | Full model for factors A, B, and C |
   | A+B+C+A*D | Main effects for A, B, and C plus the interaction of A with the covariate D |
   | A+D+D*D | Main Effect for A plus D and the square of D |
   | ~~A+B*B~~ | Not valid since B is categorical and cannot be squared |

3. Use the '|' (bar) symbol as a shorthand technique for specifying large models quickly.

   A|B = A+B+A*B

   A|B|C = A+B+C+A*B+A*C+B*C+A*B*C

   A|B C D*D = A+B+A*B+C+D*D

   A|B C|D = A+B+A*B+C+D+C*D

4. You can use parentheses for multiplication.

   (A+B)*(C+D) = A*C+A*D+B*C+B*D

   (A+B)|C = A+B+C+(A+B)*C = A+B+C+A*C+B*C

5. Use the '@' (at) symbol to limit the order of interaction terms in the model. The maximum term order can also be limited using the 'Max Term Order' function.

   A|B|C @2 = A+B+C+A*B+A*C+B*C

   A|B|D|D (Max Term Order=2) = A+B+D+A*B+A*D+B*D+D*D

### Intercept

Check this box to include the intercept in the model. Under most circumstances, you will want to include an intercept term in your model.

## Random Model (Subject Terms Only)

This section defines the random effects in the mixed model. Every term in the random model must have the Subject variable in the term. This random component can be used in specifying traditional variance component models as well as random coefficient models. Additional random components may be specified on the More Models tab. Hierarchical models with two levels of hierarchy can not be specified in the Mixed Models procedure. For example, if a study involves repeated measurements on randomly selected patients from randomly selected hospitals, only Patient or Hospital can be selected as the subject variable; and the random model can consist only of terms with the chosen variable in each term.

The purpose of this model is to define the structures of the **Z** and **G** matrices in the mixed model, as well as the random effects in the model. The **Z** matrix for random effects is comparable in function to **X** (or design) matrix for fixed effects. The **G** matrix is formed to correspond to the random effects in **Z**. For more information, see the discussion on random effects earlier in this chapter.

### Model

Specify the random component of the model here. Every term in the random model must have the Subject variable in the term. For a Random Effects model, enter the subject variable here, e.g. 'Subject'. For a Random Coefficients model, enter the subject variable and the subject variable times the time variable, e.g. 'Subject  Subject*Time'.

Try to keep this model as simple as possible.

### Groups

Specify a grouping variable here. A new set of parameters for this component will be generated for each unique value of this variable.

WARNING: because this option can quickly double or triple the number of variance parameters in the model, extreme care must be exercised when using this option.

### Covariances

If this box is checked, the G-matrix (covariance matrix) will include covariances for each pair of variance components (diagonal element of the G-matrix). If the box is not checked, all off-diagonal elements will be set to zero (the G-matrix will be diagonal).

This option is commonly checked when you are fitting a random coefficients model.

## Repeated (Time) Covariance Pattern

The repeated component is used to specify the R matrix in the mixed model. At least a diagonal pattern should always be used.

### Pattern

Specify the type of **R** (error covariance) matrix to be generated. This represents the relationship between observations from the same subject. The **R** structures that can be specified in ***NCSS*** are shown below. The usual type is the 'Diagonal' matrix.

The options are:

- **Unused**

  No repeated component is used.

- **Diagonal**

  Homogeneous

  $$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

  Heterogeneous

  $$\begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{pmatrix}$$

- **Compound Symmetry**

  Homogeneous

  $$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

  $$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \rho\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho\sigma_4\sigma_1 & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **AR(1)**

  Homogeneous

  $$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 \\ \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

  $$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \rho^3\sigma_1\sigma_4 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho^2\sigma_2\sigma_4 \\ \rho^2\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho^3\sigma_4\sigma_1 & \rho^2\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **AR(Time Diff)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho^{t_2-t_1}\sigma^2 & \rho^{t_3-t_1}\sigma^2 & \rho^{t_4-t_1}\sigma^2 \\ \rho^{t_2-t_1}\sigma^2 & \sigma^2 & \rho^{t_3-t_2}\sigma^2 & \rho^{t_4-t_2}\sigma^2 \\ \rho^{t_3-t_1}\sigma^2 & \rho^{t_3-t_2}\sigma^2 & \sigma^2 & \rho^{t_4-t_3}\sigma^2 \\ \rho^{t_4-t_1}\sigma^2 & \rho^{t_4-t_2}\sigma^2 & \rho^{t_4-t_3}\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho^{t_2-t_1}\sigma_1\sigma_2 & \rho^{t_3-t_1}\sigma_1\sigma_3 & \rho^{t_4-t_1}\sigma_1\sigma_4 \\ \rho^{t_2-t_1}\sigma_2\sigma_1 & \sigma_2^2 & \rho^{t_3-t_2}\sigma_2\sigma_3 & \rho^{t_4-t_2}\sigma_2\sigma_4 \\ \rho^{t_3-t_1}\sigma_3\sigma_1 & \rho^{t_3-t_2}\sigma_3\sigma_2 & \sigma_3^2 & \rho^{t_4-t_3}\sigma_3\sigma_4 \\ \rho^{t_4-t_1}\sigma_4\sigma_1 & \rho^{t_4-t_2}\sigma_4\sigma_2 & \rho^{t_4-t_3}\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **Toeplitz (All)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 & \rho_3\sigma^2 \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ \rho_3\sigma^2 & \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \rho_3\sigma_1\sigma_4 \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 \\ \rho_2\sigma_3\sigma_1 & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ \rho_3\sigma_4\sigma_1 & \rho_2\sigma_4\sigma_2 & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **Toeplitz(2)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & & \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \\ & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ & & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & & \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \\ & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ & & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

  Note: This is the same as Banded(2).

- **Toeplitz(3)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 & \\ \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 & \rho_1\sigma^2 \\ & \rho_2\sigma^2 & \rho_1\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \\ \rho_1\sigma_2\sigma_1 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 \\ \rho_2\sigma_3\sigma_1 & \rho_1\sigma_3\sigma_2 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 \\ & \rho_2\sigma_4\sigma_2 & \rho_1\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **Toeplitz(4) and Toeplitz(5)**

  Toeplitz(4) and Toeplitz(5) follow the same pattern as Toeplitz(2) and Toeplitz(3), but with the corresponding numbers of bands.

- **Banded(2)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & & \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \\ & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ & & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & & \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \\ & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ & & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

  Note: This is the same as Toeplitz(1).

- **Banded(3)**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 \\ \rho\sigma_3\sigma_1 & \rho\sigma_3\sigma_2 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ & \rho\sigma_4\sigma_2 & \rho\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

- **Banded(4) and Banded (5)**

  Banded(4) and Banded(5) follow the same pattern as Banded(2) and Banded(3), but with the corresponding numbers of bands.

- **Unstructured**

  Homogeneous

$$\begin{pmatrix} \sigma^2 & \rho_{12}\sigma^2 & \rho_{13}\sigma^2 & \rho_{14}\sigma^2 \\ \rho_{21}\sigma^2 & \sigma^2 & \rho_{23}\sigma^2 & \rho_{24}\sigma^2 \\ \rho_{31}\sigma^2 & \rho_{32}\sigma^2 & \sigma^2 & \rho_{34}\sigma^2 \\ \rho_{41}\sigma^2 & \rho_{42}\sigma^2 & \rho_{43}\sigma^2 & \sigma^2 \end{pmatrix}$$

  Heterogeneous

$$\begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 \\ \rho_{31}\sigma_3\sigma_1 & \rho_{32}\sigma_3\sigma_2 & \sigma_3^2 & \rho_{34}\sigma_3\sigma_4 \\ \rho_{41}\sigma_4\sigma_1 & \rho_{42}\sigma_4\sigma_2 & \rho_{43}\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

## Groups

Specify a grouping variable here. A new set of parameters for this component will be generated for each unique value of this variable.

WARNING: because this option can quickly double or triple the number of variance parameters in the model, extreme care must be exercised when using this option.

# Comparisons Tab

This panel is used to specify multiple comparisons or custom contrasts for factor variables.

## Multiple Comparisons – Default Factor Comparisons

This section allows the user to specify the default factor comparison, along with other factor comparisons.

### Comparison

The Default Comparison is used for all factors that are not specified under Factor Variable (and when Comparisons are selected under the Reports tab). For interactions, these comparisons are run for each category of the second factor. Possible choices are:

- **First versus Each**

  The multiple comparisons are each category tested against the first category. This option would be used when the first category is the control (standard) category. Note: the first is determined alphabetically.

- **2nd versus Each**

  The multiple comparisons are each category tested against the second category. This option would be used when the second category is the control (standard) category.

- **3rd versus Each**

  The multiple comparisons are each category tested against the third category. This option would be used when the third category is the control (standard) category.

- **Last versus Each**

  The multiple comparisons are each category tested against the last category. This option would be used when the last category is the control (standard) category.

- **Baseline versus Each**

  The multiple comparisons are each category tested against the baseline category. This option would be used when the baseline category is the control (standard) category. The baseline category is entered to the right.

- **Ave versus Each**

  The multiple comparisons are each category tested against the average of the other categories.

- **All Pairs**

  The multiple comparisons are each category tested against every other category.

### Baseline

Enter the level of all factor variables not specified under Factor Variable to which comparisons will be made. The Default Baseline is used only when Default Comparison is set to 'Baseline vs Each'.

The value entered here must be one of the levels of all factor variables not specified under Factor Variable. The entry is not case sensitive, and values should be entered without quotes.

## Multiple Comparisons – User-Specified Factor Comparisons

### Factor Variable

Specify settings for a particular factor variable here. All factors that are not specified here use the DEFAULT settings at the top.

Note that any variables specified here that are not specified as factors are ignored.

### Comparison

The Default Comparison is used for all factors that are not specified under Factor Variable (and when Comparisons by Design are selected under the Reports tab). For interactions, these comparisons are run for each category of the second factor. Possible choices are shown above.

### Baseline

Enter the level of the corresponding Factor Variable to which comparisons will be made. The Baseline is used only when Comparison is set to 'Baseline vs Each'. The value entered here must be one of the levels of the Factor Variable. The entry is not case sensitive and values should not be entered with quotes.

### Custom

This option specifies the weights of a comparison. It is used when the Comparison is set to 'Custom'.

NOTE: There are no numerical restrictions on these coefficients. They do not even have to sum to zero. However, this is recommended. If the coefficients do sum to zero, the comparison is called a CONTRAST. The significance tests anticipate that only one or two of these comparisons are run. If you run several, you should make some type of Bonferroni adjustment to your alpha value.

#### Specifying the Weights

When you put in your own contrasts, you must be careful that you specify the appropriate number of weights. For example, if the factor has four levels, four weights must be specified, separated by blanks or commas. Extra weights are ignored. If too few weights are specified, the missing weights are assumed to be zero.

These comparison coefficients designate weighted averages of the level-means that are to be statistically tested. The null hypothesis is that the weighted average is zero. The alternative hypothesis is that the weighted average is nonzero. The weights (comparison coefficients) are specified here in this box.

As an example, suppose you want to compare the average of the first two levels with the average of the last two levels in a six-level factor. You would enter -1 -1 0 0 1 1.

As a second example, suppose you want to compare the average of the first two levels with the average of the last three levels in a six-level factor. The custom contrast would be -3 -3 0 2 2 2.

Note that in each example, weights were used that sum to zero. Ones were not used in the second example because the result would not sum to zero.

## Comparisons Using a User-Specified Contrast (L) Matrix

### L-Matrix Variables

Specify one or more variables (columns) containing a contrast matrix that you want to test. This allows you to test any contrast you want. The layout of the contrast matrix is identical to the layout that is displayed when the L-Matrices are output. Hence, we suggest you first run an analysis, output the L-Matrices, and then use these output L-matrices as a template.

Note: Only one L-matrix can be entered at a time. If you want create multiple tests, you will have to do multiple runs.

# Covariates Tab

This panel is used to define the covariate values at which means and comparisons of other factors will be computed.

## Covariate Variable Settings – Default Factor Comparisons

This section allows the user to specify the default covariate value(s) at which means of other factors will be computed.

### Compute Means at these Values

This is the value (or values) used for each covariate that is not specified under Covariate Variable below. Means and comparisons are computed at this value.

## Covariate Variable Settings – User-Specified Covariate Settings

This section allows the user to specify the covariate value(s) at which means and comparisons of other factors will be computed.

### Covariate Variable

Specify a Covariate Variable for which means and comparisons will be computed at a specific value. Covariates specified here must be in the Covariate Variables list of the Variables tab.

### Compute Means at these Values

Specify one or more values of the corresponding Covariate Variable at which means and planned comparisons will be calculated. A separate analysis is calculated for each value entered here. When more than one Covariate Variable is specified, a separate analysis is carried out for each combination of covariate values.

# Reports Tab

The following options control which plots and reports are displayed.

## Select Reports

### Run Summary Report

Check this box to obtain a summary of the likelihood type, the model, the iterations, the resulting likelihood/AIC, and run time.

### Variance Estimates Report

Check this box to obtain estimates of random and repeated components of the model.

### Hypothesis Tests Report

Check this box to obtain F-Tests for all terms in the Fixed (Means) Specification (see Variables tab).

### L-Matrices – Terms Report

Check this box to obtain L matrices for each term in the model. Each L matrix describes the linear combination of the betas that is used to test the corresponding term in the model.

Caution: Selecting this option can generate a very large amount of output, as the L matrices can be very numerous and lengthy.

### Comparisons by Fixed Effects Report

Check this box to obtain planned comparison tests, comparing levels of the fixed effects. Details of the comparisons to be made are specified under the Comparisons and Covariates tabs. When more than one covariate value is specified under the Covariates tab, the comparisons are grouped such that for each fixed effect, comparisons for all covariate(s) values are displayed.

Compare to Comparisons by Covariate Values.

### Comparisons by Covariate Values Report

Check this box to obtain planned comparison tests, comparing levels of the fixed effects. Details of the comparisons to be made are specified under the Comparisons and Covariates tabs. When more than one covariate value is specified under the Covariates tab, the comparisons are grouped such that for each value of the covariate(s), a new set of comparisons is displayed.

Compare to Comparisons by Fixed Effects.

### L-Matrices – Comparisons Report

Check this box to obtain L matrices for each planned comparison. Each L matrix describes the linear combination of the betas that is used to test the corresponding comparison.

Caution: Selecting this option can generate a very large amount of output, as the L matrices can be very numerous and lengthy.

### Means by Fixed Effects Report

Check this box to obtain means and confidence limits for each fixed effect level. When more than one covariate value is specified under the Covariates tab, the means are grouped such that for each fixed effect, means for all covariate(s) values are displayed.

Compare to Means by Covariate Values.

### Means by Covariate Values Report

Check this box to obtain means and confidence limits for each fixed effect level. When more than one covariate value is specified under the Covariates tab, the means are grouped such that for each value of the covariate(s), a new set of means is displayed.

Compare to Means by Fixed Effects.

### L-Matrices – LS Means Report

Check this box to obtain L matrices for each least squares mean (of the fixed effects). Each L matrix describes the linear combination of the betas that is used to generate the least squares mean.

Caution: Selecting this option can generate a very large amount of output, as the L matrices can be very numerous and lengthy.

### Fixed Effects Solution Report

Check this box to obtain estimates, P-values and confidence limits of the fixed effects and covariates (betas).

### Asymptotic VC Matrix Report

Check this box to obtain the asymptotic variance-covariance matrix of the random (and repeated) components of the model..

### Vi Matrices (1st 3 Subjects) Report

Check this box to display the Vi matrices of the first three subjects.

### Hessian Matrix Report

Check this box to obtain the Hessian matrix. The Hessian matrix is directly associated with the variance-covariance matrix of the random (and repeated) components of the model.

## Select Plots

### Means Plots

Check this box to obtain plots of means for each fixed effects term of the model. Details of the appearance of the plots are specified under the Means Plots and Symbols tabs.

### Subject Plots

Check this box to obtain plots of the repeated values for each subject. Plots comparing main effects for each subject are also given. The repeated values for each subject are ordered according to the order the values appear in the data set. Details of the appearance of the plots are specified under the Subject Plots and Symbols tabs.

## Report Options

### Alpha

Specify the alpha value (significance level) used for F-tests, T-tests, and confidence intervals. Alpha is the probability of rejecting the null hypothesis of equal means when it is actually true. Usually, an alpha of .05 is used. Typical choices for alpha range between .001 and .200.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, while the double precision will display thirteen-place accuracy.

### Show Notes

Indicate whether to show the notes at the end of reports. Although these notes are helpful at first, they may tend to clutter the output. This option lets you omit them.

## Report Options – Decimal Places

### Effects/Betas ... Covariates

Specify the number of digits after the decimal point to display on the output of values of this type. Note that this option in no way influences the accuracy with which the calculations are done.

Enter 'General' to display all digits available. The number of digits displayed by this option is controlled by whether the PRECISION option is SINGLE or DOUBLE.

# Maximization Tab

This tab controls the Newton-Raphson, Fisher-Scoring, and Differential Evolution likelihood-maximization algorithms.

## Newton-Raphson / Fisher-Scoring Options

### Max Fisher Scoring Iterations

This is the maximum number of Fisher Scoring iterations that occur in the maximum likelihood finding process. When Solution Method (Variables tab) is set to 'Newton-Raphson', up to this number of Fisher Scoring iterations occur before beginning Newton-Raphson iterations.

### Max Newton-Raphson Iterations

This is the maximum number of Newton-Raphson iterations that occur in the maximum likelihood finding process. When Solution Method (Variables tab) is set to 'Newton-Raphson', Fisher-scoring iterations occur before beginning Newton-Raphson iterations.

### Lambda

Each parameter's change is multiplied by this value at each iteration. Usually, this value can be set to one. However, it may be necessary to set this value to 0.5 to implement step-halving: a process that is necessary when the Newton-Raphson diverges.

Note: this parameter only used by the Fisher-Scoring and Newton-Raphson methods.

### Convergence Criterion

This procedure uses relative Hessian convergence (or the Relative Offset Orthogonality Convergence Criterion) as described by Bates and Watts (1981).

Recommended: The default value, 1E-8, will be adequate for many problems. When the routine fails to converge, try increasing the value to 1E-6.

## Differential Evolution Options

### Crossover Rate

This value controls the amount of movement of the differential evolution algorithm toward the current best. Larger values accelerate movement toward the current best, but reduce the chance of locating the global maximum. Smaller values improve the chances of finding the global, rather than a local, solution, but increase the number of iterations until convergence.

RANGE: Usually, a value between .5 and 1.0 is used.

RECOMMENDED: 0.9.

### Mutation Rate

This value sets the mutation rate of the search algorithm. This is the probability that a parameter is set to a random value within the parameter space. It keeps the algorithm from stalling on a local maximum.

RANGE: Values between 0 and 1 are allowed.

RECOMMENDED: 0.9 for random coefficients (complex) models or 0.5 for random effects (simple) models.

### Minimum Relative Change

This parameter controls the convergence of the likelihood maximizer. When the relative change in the likelihoods from one generation to the next is less than this amount, the algorithm concludes that it has converged. The relative change is |L(g+1) - L(g)| / L(g) where L(g) is absolute value of the likelihood at generation 'g'. Note that the algorithm also terminates if the Maximum Generations are reached or if the number of individuals that are replaced in a generation is zero. The value 0.00000000001 (ten zeros) seems to work well in practice. Set this value to zero to ignore this convergence criterion.

### Solutions/Iteration

This is the number of trial points (solution sets) that are used by the differential evolution algorithm during each iteration. In the terminology of differential evolution, this is the population size.

RECOMMENDED: A value between 15 and 25 is recommended. More points may dramatically increase the running time. Fewer points may not allow the algorithm to converge.

### Max Iterations

Specify the maximum number of differential evolution iterations used by the differential evolution algorithm. A value between 100 and 200 is usually adequate. For large datasets, i.e., number of rows greater than 1000, you may want to reduce this number.

## Other Options

### Max Retries

Specify the maximum number of retries to occur. During the maximum likelihood search process, the search may lead to an impossible combination of variance-covariance parameters (as defined by a matrix of variance-covariance parameters that is not positive definite). When such a combination arises, the search algorithm will begin again. Max Retries is the maximum number of times the process will re-start to avoid such combinations.

### Zero (Algorithm Rounding)

This cutoff value is used by the least-squares algorithm to lessen the influence of rounding error. Values lower than this are reset to zero. If unexpected results are obtained, try using a smaller value, such as 1E-32. Note that 1E-5 is an abbreviation for the number 0.00001.

RECOMMENDED: 1E-10 or 1E-12.

RANGE: 1E-3 to 1E-40.

### Variance Zero

When an estimated variance component (diagonal element) is less than this value, the variance is assumed to be zero and all reporting is terminated since the algorithm has not converged properly.

To correct this problem, remove the corresponding term from the Random Factors Model or simplify the Repeated Variance Pattern. Since the parameter is zero, why would you want to keep it?

RECOMMENDED: 1E-6 or 1E-8.

RANGE: 1E-3 to 1E-40.

### Correlation Zero

When an estimated correlation (off-diagonal element) is less than this value, the correlation is assumed to be zero and all reporting is terminated since the algorithm has not converged properly.

To correct this problem, remove the corresponding term from the Random Factors Model or simplify the Repeated Variance Pattern. Since the parameter is zero, why would you want to keep it?

RECOMMENDED: 1E-6 or 1E-8.

RANGE: 1E-3 to 1E-40.

## More Models Tab

This tab allows the user to specify random and repeated model components in addition to those specified on the Variables tab.

### More Random Models (Subject Only)

#### Model

Specify the random (subject) component of the model here. For a Random Effects model, enter the subject variable here, e.g. 'Subject'. For a Random Coefficients model, enter the subject variable and the subject variable times the time variable, e.g. 'Subject  Subject*Time'.

Every term of a random model must include the Subject variable as part of the term.

In general, random models should be as simple as possible.

#### Groups

Specify a grouping variable here. A new set of parameters for this component will be generated for each unique value of this variable.

WARNING: because this option can quickly double or triple the number of variance parameters in the model, extreme care must be exercised when using this option.

### Covariances

If this box is checked, the G-matrix (covariance matrix) will include covariances for each pair of variance components (diagonal element of the G-matrix). If the box is not checked, all off-diagonal elements will be set to zero (the G-matrix will be diagonal).

This option is commonly checked when you are fitting a random coefficient model.

## More Repeated Covariance Patterns

### Pattern

Specify the type of R (error covariance) matrix to be generated. The usual type is the 'Diagonal' matrix.

### Groups

Specify a grouping variable here. A new set of parameters for this component will be generated for each unique value of this variable.

WARNING: because this option can quickly double or triple the number of variance parameters in the model, extreme care must be exercised when using this option.

# Means Plot Tab

These options specify the plots of group means.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Y Scaling

Specify the method for calculating the minimum and maximum along the vertical axis. *Separately* means that each plot is scaled independently. *Uniform* means that all plots use the overall minimum and maximum of the data. This option is ignored if a minimum or maximum is specified.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Means Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect All Points

Check this box to connect the points on the plot with a line. The line is drawn sequentially, using the order of the rows on the database.

### Symbol Radius %

Reduce (or increase) the radius of all plot symbols by this percentage amount. This option was added so that you can quickly resize all of the plot symbols with a single change. The value must be a number between 1 and 1000.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, *{G}* is replaced by the appropriate factor name.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Plot Subtitle

This is the text of the plot subtitle. This is usually used to display covariate values on the plot.

# Subject Plots Tab

These options specify the subject plots.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Subject Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

## Plot Settings – Plot Grouped by Subject

These options control the settings of the first subject plot, which is grouped by subject.

### Connect All Points

Check this box to connect the points on the plot with a line. The line is drawn sequentially, using the order of the rows on the database.

### Symbol Radius %

Reduce (or increase) the radius of all plot symbols by this percentage amount. This option was added so that you can quickly resize all of the plot symbols with a single change. The value must be a number between 1 and 1000.

## Plot Settings – Plots Grouped by Factor

These options control the settings of the other subject plots, which are grouped by factor.

### Connect All Points

Check this box to connect the points on the plot with a line. The line is drawn sequentially, using the order of the rows on the database.

### Symbol Radius %

Reduce (or increase) the radius of all plot symbols by this percentage amount. This option was added so that you can quickly resize all of the plot symbols with a single change. The value must be a number between 1 and 1000.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, *{G}* is replaced by the appropriate factor name.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* and *{X}* are replaced by appropriate names. The character *{G}* is replaced by the appropriate factor name. Press the button on the right of the field to specify the font of the text.

# Symbols Tab

These options specify the symbols used in the plots.

## Plotting Symbols

### Group (1-15)

The symbols used to represent the levels of a factor on the means plots. Group 1 represents the first level, Group 2 represents the second level, and so on.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Longitudinal Design (One Between-Subject Factor, One Within-Subject Factor, One Covariate)

This example has two purposes:

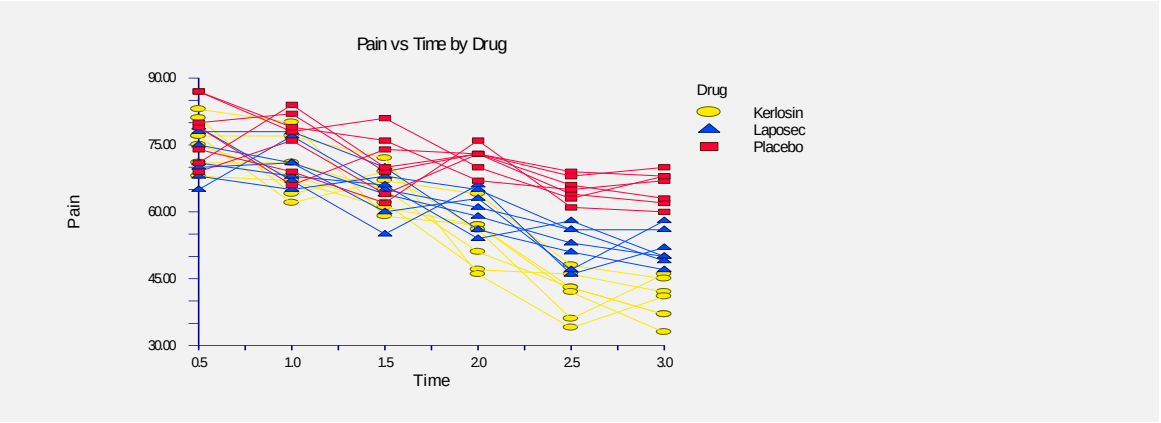
1. Acquaint the reader with the output for all output options. In only this example, each heading of each section of the output is described in detail.
2. Describe a typical analysis of a longitudinal design. A portion of this example involves the comparison of options for the Repeated Variance Pattern. There is some discussion as the output is presented and annotated, with a fuller discussion of model refinement and covariance options at the end of this example.

In a longitudinal design, subjects are measured more than once, usually over time. This example presents the analysis of a longitudinal design in which there is one between-subjects factor, one within-subjects factor (Time), and a covariate. Two drugs (Kerlosin and Laposec) are compared to a placebo for their effectiveness in reducing pain following a surgical eye procedure. A standard pain measurement for each patient is measured at 30 minute intervals following surgery and administration of the drug (or placebo). Six measurements, with the last at Time = 3 hours, are made for each of the 21 patients (7 per group). A blood pressure measurement of each individual at the time of pain measurement is measured as a covariate. The researchers wish to compare the drugs at the covariate value of 140.

**PAIN Dataset**

| Drug | Patient | Time | Cov | Pain |
|---|---|---|---|---|
| Kerlosin | 1 | 0.5 | 125 | 68 |
| Kerlosin | 1 | 1 | 196 | 67 |
| Kerlosin | 1 | 1.5 | 189 | 61 |
| Kerlosin | 1 | 2 | 135 | 57 |
| Kerlosin | 1 | 2.5 | 128 | 43 |
| Kerlosin | 1 | 3 | 151 | 37 |
| Kerlosin | 2 | 0.5 | 215 | 75 |
| Kerlosin | 2 | 1 | 151 | 68 |
| Kerlosin | 2 | 1.5 | 191 | 62 |
| Kerlosin | 2 | 2 | 212 | 47 |
| Kerlosin | 2 | 2.5 | 127 | 46 |
| Kerlosin | 2 | 3 | 133 | 42 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| Placebo | 21 | 2 | 129 | 73 |
| Placebo | 21 | 2.5 | 216 | 68 |
| Placebo | 21 | 3 | 158 | 70 |

The following plot shows the relationship among all variables except the covariate.




To run the analysis using the Mixed Models procedure, you can enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 1** from the Template tab of the Mixed Models window.

**1 Open the PAIN dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **PAIN.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Pain** from the list of variables and then click **Ok**. 'Pain' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Patient** from the list of variables and then click **Ok**. 'Patient' will appear in the Subject Variable box.
- Select **Time** for the **Time Variable** text box.
- Select **Drug**, **Time** for the **Factor (Categorical) Variables** text box.
- Select **Cov** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Drug Time Drug*Time Cov Drug*Cov Time*Cov Drug*Time*Cov** under **Model** for the **Fixed Effects Model**.
- Enter **Patient** under **Model** for the **Random Model (Subject Terms Only)**.

**5 Specify the likelihood options.**

- Leave the **Likelihood** as **REML** and the **Solution Method** as **Newton-Raphson**.

**6 Specify the comparisons.**

- On the Mixed Models window, select the **Comparisons tab**.
- Select **Drug** as the **first Factor Variable**. Select **All Pairs** for the **Comparison**.
- Select **Time** as the **second Factor Variable**. Select **Baseline vs Each** for the **Comparison**. Enter **0.5** for **Baseline**.

**7 Specify the covariate.**

- On the Mixed Models window, select the **Covariates tab**.
- Select **Cov** as a **Covariate Variable**. Enter **140** for **Compute Means at these values**.

**8 Specify the reports.**

- On the Mixed Models window, select the **Reports tab**.
- Check all report and plot checkboxes except **L Matrices – Comparisons** and **L Matrices – LS Means**.

**9 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Run Summary Section

**Run Summary Section**

| Parameter | Value |
|---|---|
| Likelihood Type | Restricted Maximum Likelihood |
| Fixed Model | DRUG+TIME+COV+DRUG*TIME+DRUG*COV+TIME*COV+ DRUG*TIME*COV |
| Random Model | PATIENT |
| Repeated Model | Diagonal |
| Number of Rows | 126 |
| Number of Subjects | 21 |
| Solution Type | Newton-Raphson |
| Fisher Iterations | 4 of a possible 10 |
| Newton Iterations | 1 of a possible 40 |
| Max Retries | 10 |
| Lambda | 1 |
| Log Likelihood | -369.16 |
| -2 Log Likelihood | 738.31 |
| AIC (Smaller Better) | 742.31 |
| Convergence | Normal |
| Run Time (Seconds) | 8.515625 |

This section provides a summary of the model and the iterations toward the maximum log likelihood.

### Likelihood Type

This value indicates that restricted maximum likelihood was used rather than maximum likelihood.

### Fixed Model

The model shown is that entered as the Fixed Factors Model of the Variables tab. The model includes fixed terms and covariates.

### Random Model

The model shown is that entered as the Random Factors Model of the Variables tab.

### Repeated Model

The pattern shown is that entered as the Repeated (Time) Variance Pattern of the Variables tab.

### Number of Rows

The number of rows processed from the database.

### Number of Subjects

The number of unique subjects from the database.

### Solution Type

The solution type is method used for finding the maximum (restricted) maximum likelihood solution. Newton-Raphson is the recommended method.

### Fisher Iterations

Some Fisher-Scoring iterations are used as part of the Newton-Raphson algorithm. The ‘4 of a possible 10’ means four Fisher-Scoring iterations were used, while ten was the maximum that were allowed (as specified on the Maximization tab).

### Newton Iterations

The ‘1 of a possible 40’ means one Newton-Raphson iteration was used, while forty was the maximum allowed (as specified on the Maximization tab).

### Max Retries

The maximum number of times that lambda was changed and new variance-covariance parameters found during an iteration was ten. If the values of the parameters result in a negative variance, lambda is divided by two and new parameters are generated. This process continues until a positive variance occurs or until Max Retries is reached.

### Lambda

Lambda is a parameter used in the Newton-Raphson process to specify the amount of change in parameter estimates between iterations. One is generally an appropriate selection. When convergence problems occur, reset this to 0.5.

If the values of the parameters result in a negative variance, lambda is divided by two and new parameters are generated. This process continues until a positive variance occurs or until Max Retries is reached.

### Log Likelihood

This is the log of the likelihood of the data given the variance-covariance parameter estimates. When a maximum is reached, the algorithm converges.

### -2 Log Likelihood

This is minus 2 times the log of the likelihood. When a minimum is reached, the algorithm converges.

### AIC

The Akaike Information Criterion is used for comparing covariance structures in models. It gives a penalty for increasing the number of covariance parameters in the model.

### Convergence

'Normal' convergence indicates that convergence was reached before the limit.

### Run Time (Seconds)

The run time is the amount of time used to solve the problem and generate the output.

## Random Component Parameter Estimates (G Matrix)

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 1.6343 | Patient |

This section gives the random component estimates according to the Random Factors Model specifications of the Variables tab.

### Component Number

A number is assigned to each random component. The first component is the one specified on the variables tab. Components 2-5 are specified on the More Models tab.

### Parameter Number

When the random component model results in more than one parameter for the component, the parameter number identifies parameters within the component.

### Estimated Value

The estimated value 1.6343 is the estimated patient variance component.

### Model Term

Patient is the name of the random term being estimated.

## Repeated Component Parameter Estimates (R Matrix)

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 23.5867 | Diagonal (Variance) |

This section gives the repeated component estimates according to the Repeated Variance Pattern specifications of the Variables tab.

### Component Number

A number is assigned to each repeated component. The first component is the one specified on the variables tab. Components 2-5 are specified on the More Models tab.

### Parameter Number

When the repeated pattern results in more than one parameter for the component, the parameter number identifies parameters within the component.

### Estimated Value

The estimated value 23.5867 is the estimated residual (error) variance.

### Parameter Type

The parameter type describes the structure of the R matrix that is estimated, and is specified by the Repeated Component Pattern of the Variables tab.

## Term-by-Term Hypothesis Test Results

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Drug | 1.82 | 2 | 89.0 | 0.1677 |
| Time | 0.98 | 5 | 88.4 | 0.4358 |
| Cov | 3.30 | 1 | 87.1 | 0.0726 |
| Drug*Time | 0.86 | 10 | 87.0 | 0.5708 |
| Drug*Cov | 0.77 | 2 | 86.8 | 0.4662 |
| Time*Cov | 1.22 | 5 | 88.5 | 0.3078 |
| Drug*Time*Cov | 1.07 | 10 | 87.0 | 0.3947 |

These F-Values test Type-III (adjusted last) hypotheses.

This section contains a F-test for each component of the Fixed Component Model according to the methods described by Kenward and Roger (1997).

### Model Term

This is the name of the term in the model.

### F-Value

The F-Value corresponds to the L matrix used for testing this term in the model. The F-Value is based on the F approximation described in Kenward and Roger (1997).

### Num DF

This is the numerator degrees of freedom for the corresponding term.

### Denom DF

This is the approximate denominator degrees of freedom for this comparison as described in Kenward and Roger (1997).

### Prob Level

The Probability Level (or P-value) gives the strength of evidence (smaller Prob Level implies more evidence) that a term in the model has differences among its levels, or a slope different from zero in the case of covariate. It is the probability of obtaining the corresponding F-Value (or greater) if the null hypothesis of equal means (or no slope) is true.

# Individual Comparison Hypothesis Test Results

**Individual Comparison Hypothesis Test Results**
**Covariates: Cov=140.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Drug | | 43.18 | 2 | 32.8 | 0.0000 | |
| Drug: Kerlosin - Laposec | -3.47 | 4.23 | 1 | 37.7 | 0.0467 | 0.1402 [3] |
| Drug: Kerlosin - Placebo | -13.60 | 78.75 | 1 | 28.9 | 0.0000 | 0.0000 [3] |
| Drug: Laposec - Placebo | -10.13 | 39.47 | 1 | 33.8 | 0.0000 | 0.0000 [3] |
| Time | | 46.51 | 5 | 82.3 | 0.0000 | |
| Time: 0.5 - 1 | 2.81 | 1.36 | 1 | 87.0 | 0.2467 | 1.0000 [5] |
| Time: 0.5 - 1.5 | 8.19 | 20.23 | 1 | 82.7 | 0.0000 | 0.0001 [5] |
| Time: 0.5 - 2 | 11.30 | 29.22 | 1 | 79.6 | 0.0000 | 0.0000 [5] |
| Time: 0.5 - 2.5 | 21.26 | 122.12 | 1 | 83.6 | 0.0000 | 0.0000 [5] |
| Time: 0.5 - 3 | 22.26 | 152.66 | 1 | 81.0 | 0.0000 | 0.0000 [5] |
| Drug*Time | | 5.38 | 10 | 81.1 | 0.0000 | |
| Drug = Kerlosin, Time: 0.5 - 1 | 7.04 | 3.70 | 1 | 86.3 | 0.0578 | 0.8674 [15] |
| Drug = Kerlosin, Time: 0.5 - 1.5 | 10.05 | 9.45 | 1 | 84.9 | 0.0028 | 0.0426 [15] |
| Drug = Kerlosin, Time: 0.5 - 2 | 19.70 | 19.57 | 1 | 77.8 | 0.0000 | 0.0005 [15] |
| Drug = Kerlosin, Time: 0.5 - 2.5 | 34.29 | 118.81 | 1 | 80.8 | 0.0000 | 0.0000 [15] |

(report continues)

This section shows the F-tests for comparisons of the levels of the fixed terms of the model according to the methods described by Kenward and Roger (1997). The individual comparisons are grouped into subsets of the fixed model terms.

### Comparison/Covariate(s)

This is the comparison being made. The first line is ‘Drug’. On this line, the levels of drug are compared when the covariate is equal to 140. The second line is ‘Drug: Placebo – Kerlosin’. On this line, Kerlosin is compared to Placebo when the covariate is equal to 140.

### Comparison Mean Difference

This is the difference in the least squares means for each comparison.

### F-Value

The F-Value corresponds to the L matrix used for testing this comparison. The F-Value is based on the F approximation described in Kenward and Roger (1997).

### Num DF

This is the numerator degrees of freedom for this comparison.

### Denom DF

This is the approximate denominator degrees of freedom for this comparison as described in Kenward and Roger (1997).

### Raw Prob Level

The Raw Probability Level (or Raw P-value) gives the strength of evidence for a single comparison, unadjusted for multiple testing. It is the single test probability of obtaining the corresponding difference if the null hypothesis of equal means is true.

### Bonferroni Prob Level

The Bonferroni Prob Level is adjusted for multiple tests. The number of tests adjusted for is enclosed in brackets following each Bonferroni Prob Level. For example, 0.8674 [15] signifies that the probability the means are equal, given the data, is 0.8674, after adjusting for 15 tests.

# Least Squares (Adjusted) Means

**Least Squares (Adjusted) Means**
**Covariates: Cov=140.00**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 64.11 | 0.66 | 62.78 | 65.45 | 33.5 |
| **Drug** | | | | | |
| Kerlosin | 58.43 | 1.14 | 56.11 | 60.74 | 32.9 |
| Laposec | 61.89 | 1.24 | 59.38 | 64.40 | 42.3 |
| Placebo | 72.02 | 1.03 | 69.91 | 74.14 | 24.8 |
| **Time** | | | | | |
| 0.5 | 75.08 | 1.39 | 72.32 | 77.85 | 89.8 |
| 1 | 72.27 | 1.99 | 68.32 | 76.22 | 89.9 |
| 1.5 | 66.89 | 1.22 | 64.46 | 69.32 | 89.3 |
| 2 | 63.79 | 1.63 | 60.55 | 67.02 | 90.0 |
| 2.5 | 53.82 | 1.37 | 51.09 | 56.55 | 89.7 |
| 3 | 52.82 | 1.20 | 50.43 | 55.22 | 89.2 |
| **Drug*Time** | | | | | |
| Kerlosin, 0.5 | 76.07 | 2.60 | 70.90 | 81.23 | 89.8 |
| Kerlosin, 1 | 69.02 | 2.64 | 63.78 | 74.26 | 90.0 |
| Kerlosin, 1.5 | 66.02 | 2.05 | 61.94 | 70.10 | 89.2 |
| Kerlosin, 2 | 56.36 | 3.78 | 48.85 | 63.88 | 89.9 |
| Kerlosin, 2.5 | 41.78 | 1.90 | 38.00 | 45.55 | 88.6 |
| Kerlosin, 3 | 41.30 | 1.99 | 37.34 | 45.27 | 89.0 |
| Laposec, 0.5 | 69.44 | 2.37 | 64.73 | 74.15 | 89.8 |
| Laposec, 1 | 71.92 | 5.00 | 61.99 | 81.85 | 89.5 |
| Laposec, 1.5 | 62.50 | 2.18 | 58.17 | 66.82 | 89.4 |
| Laposec, 2 | 62.48 | 2.28 | 57.96 | 67.01 | 89.8 |
| Laposec, 2.5 | 53.42 | 1.97 | 49.50 | 57.34 | 88.9 |
| Laposec, 3 | 51.59 | 2.21 | 47.20 | 55.98 | 89.6 |
| Placebo, 0.5 | 79.74 | 2.25 | 75.27 | 84.22 | 89.6 |
| Placebo, 1 | 75.87 | 1.91 | 72.08 | 79.67 | 88.6 |
| Placebo, 1.5 | 72.16 | 2.13 | 67.93 | 76.39 | 89.3 |
| Placebo, 2 | 72.52 | 2.09 | 68.37 | 76.67 | 89.3 |
| Placebo, 2.5 | 66.27 | 3.08 | 60.14 | 72.39 | 90.0 |
| Placebo, 3 | 65.58 | 2.05 | 61.50 | 69.65 | 89.1 |

This section gives the adjusted means for the levels of each fixed factor when Cov = 140.

### Name

This is the level of the fixed term that is estimated on the line.

### Mean

The mean is the estimated least squares (adjusted or marginal) mean at the specified value of the covariate.

### Standard Error of Mean

This is the standard error of the mean.

### 95.0% Lower (Upper) Conf. Limit for Mean

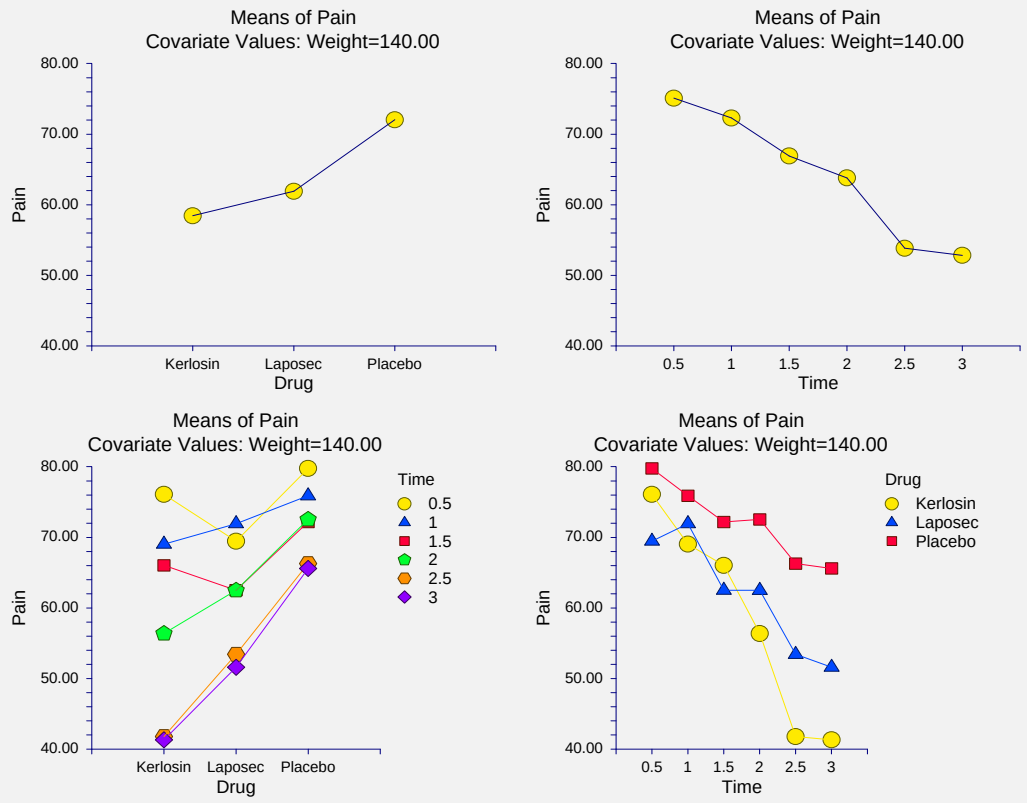
These limits give a 95% confidence interval for the mean.

### DF

The degrees of freedom used for the confidence limits are calculated using the method of Kenward and Roger (1997).

# Means Plots

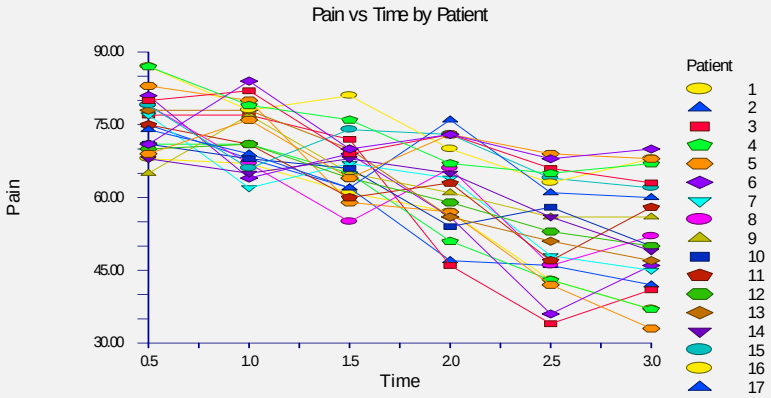
**Means Plots**

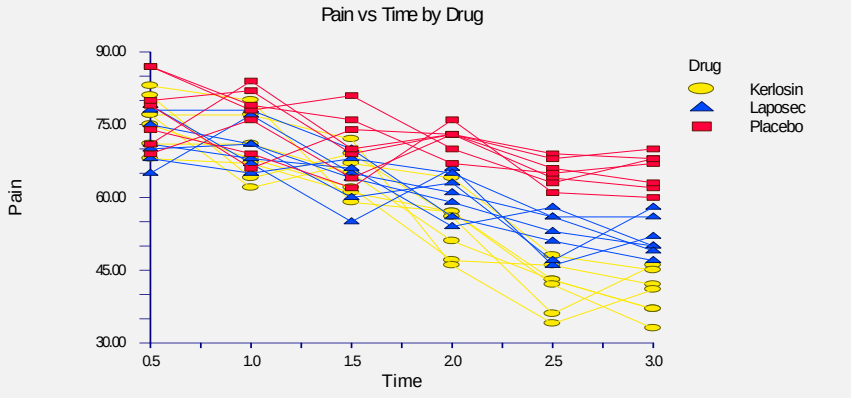


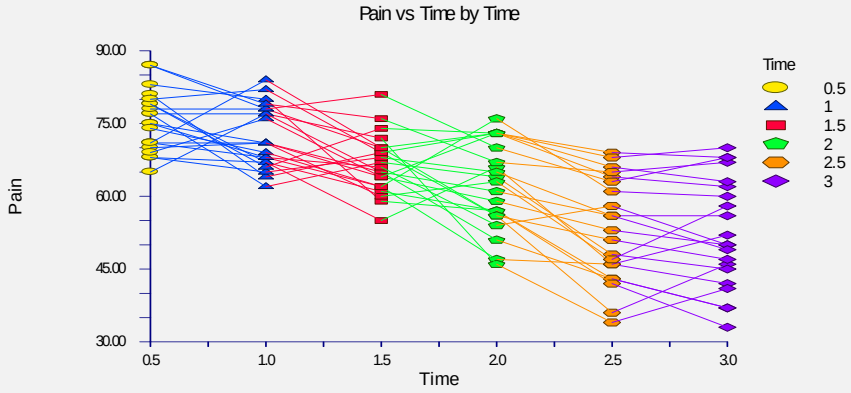

These plots show the means broken up into the categories of the fixed effects of the model. Some general trends that can be seen are those of pain decreasing with time and lower pain for the two drugs after two hours.

# Subject Plots

**Subject Plots**






Each set of connected dots of the Subject plots show the repeated measurements on the same subject. The second plot is perhaps the most telling, as it shows a separation of pain among drugs after 2 hours.

## Solution for Fixed Effects

**Solution for Fixed Effects**

| Effect No. | Effect Name | Effect Estimate (Beta) | Effect Standard Error | Prob Level | 95.0% Lower Conf. Limit of Beta | 95.0% Upper Conf. Limit of Beta | DF |
|---|---|---|---|---|---|---|---|
| 1 | Intercept | 66.8296 | 7.4693 | 0.0000 | 51.9900 | 81.6692 | 89.8 |
| 2 | (Drug="Kerlosin") | -9.4849 | 11.9162 | 0.4282 | -33.1595 | 14.1898 | 89.7 |
| 3 | (Drug="Laposec") | -16.5164 | 14.6176 | 0.2615 | -45.5591 | 12.5264 | 89.5 |
| 4 | (Drug="Placebo") | 0.0000 | 0.0000 | | | | |
| 5 | (Time=0.5) | 23.0382 | 12.0137 | 0.0584 | -0.8336 | 46.9099 | 88.8 |
| 6 | (Time=1) | -4.9520 | 10.5394 | 0.6396 | -25.9029 | 15.9988 | 86.2 |
| 7 | (Time=1.5) | 16.5033 | 12.3512 | 0.1850 | -8.0451 | 41.0518 | 87.2 |
| 8 | (Time=2) | 10.8739 | 14.7800 | 0.4640 | -18.5236 | 40.2714 | 82.9 |
| 9 | (Time=2.5) | 3.0828 | 12.5528 | 0.8066 | -21.8933 | 28.0589 | 81.0 |
| 10 | (Time=3) | 0.0000 | 0.0000 | | | | |
| 11 | Cov | -0.0089 | 0.0461 | 0.8467 | -0.1004 | 0.0826 | 89.6 |

(report continues)

This section shows the model estimates for all the model terms (betas).

### Effect No.

This number identifies the effect of the line.

### Effect Name

The Effect Name is the level of the fixed effect that is examine on the line.

### Effect Estimate (Beta)

The Effect Estimate is the beta-coefficient for this effect of the model. For main effects terms the number of effects per term is the number of levels minus one. An effect estimate of zero is given for the last effect(s) of each term. There may be several zero estimates for effects of interaction terms.

### Effect Standard Error

This is the standard error for the corresponding effect.

### Prob Level

The Prob Level tests whether the effect is zero.

### 95.0% Lower (Upper) Conf. Limit of Beta

These limits give a 95% confidence interval for the effect.

### DF

The degrees of freedom used for the confidence limits and hypothesis tests are calculated using the method of Kenward and Roger (1997).

## Asymptotic Variance-Covariance Matrix of Variance Estimates

**Asymptotic Variance-Covariance Matrix of Variance Estimates**

| Parm | G(1,1) | R(1,1) |
|---|---|---|
| G(1,1) | 4.5645 | -2.6362 |
| R(1,1) | -2.6362 | 15.0707 |

This section gives the asymptotic variance-covariance matrix of the variance components of the model. Here, the variance of the Patient variance component is 4.5645. The variance of the residual variance is 15.0707.

### Parm

Parm is the heading for both the row variance parameters and column variance parameters.

### G(1,1)

The two elements of G(1,1) refer to the component number and parameter number of the covariance parameter in G.

### R(1,1)

The two elements of R(1,1) refer to the component number and parameter number of the covariance parameter in R.

## Estimated Vi Matrix of Subject = X

**Estimated Vi Matrix of Subject = 1**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 2 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 3 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 |
| 4 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 |
| 5 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 |
| 6 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 |

**Estimated Vi Matrix of Subject = 2**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 2 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 3 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 |
| 4 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 |
| 5 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 |
| 6 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 |

**Estimated Vi Matrix of Subject = 3**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 2 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 | 1.6343 |
| 3 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 | 1.6343 |
| 4 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 | 1.6343 |
| 5 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 | 1.6343 |
| 6 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 1.6343 | 25.2210 |

This section gives the estimated variance-covariance matrix for each of the first three subjects.

### 1 – 6

Each of the 6 levels shown here represents one of the time values. That is 1 is for 0.5 hours, 2 is for 1 hour, 3 is for 1.5 hours, and so on. The number 25.2210 is calculated by adding the two variance estimates together, 1.6343 + 23.5867 = 25.2210.

## Hessian Matrix of Variance Estimates

**Hessian Matrix of Variance Estimates**

| Parm | G(1,1) | R(1,1) |
|---|---|---|
| G(1,1) | 0.2437 | 0.0426 |
| R(1,1) | 0.0426 | 0.0738 |

The Hessian Matrix is directly related to the asymptotic variance-covariance matrix of the variance estimates.

### Parm

Parm is the heading for both the row variance parameters and column variance parameters.

### G(1,1)

The two elements of G(1,1) refer to the component number and parameter number of the covariance parameter in G.

### R(1,1)

The two elements of R(1,1) refer to the component number and parameter number of the covariance parameter in R.

## L Matrices

**L Matrix for Drug**

| No. | Effect | Drug | Time | L1 | L2 |
|---|---|---|---|---|---|
| 1 | Intercept | | | | |
| 2 | Drug | Kerlosin | | 1.0000 | 1.0000 |
| 3 | Drug | Laposec | | -1.0000 | |
| 4 | Drug | Placebo | | | -1.0000 |
| 5 | Time | | 0.5 | | |
| 6 | Time | | 1 | | |
| 7 | Time | | 1.5 | | |
| 8 | Time | | 2 | | |
| 9 | Time | | 2.5 | | |
| 10 | Time | | 3 | | |
| 11 | Cov | | | | |
| 12 | Drug*Time | Kerlosin | 0.5 | 0.1667 | 0.1667 |
| 13 | Drug*Time | Kerlosin | 1 | 0.1667 | 0.1667 |
| 14 | Drug*Time | Kerlosin | 1.5 | 0.1667 | 0.1667 |
| 15 | Drug*Time | Kerlosin | 2 | 0.1667 | 0.1667 |
| 16 | Drug*Time | Kerlosin | 2.5 | 0.1667 | 0.1667 |
| 17 | Drug*Time | Kerlosin | 3 | 0.1667 | 0.1667 |
| 18 | Drug*Time | Laposec | 0.5 | -0.1667 | |
| 19 | Drug*Time | Laposec | 1 | -0.1667 | |
| 20 | Drug*Time | Laposec | 1.5 | -0.1667 | |
| 21 | Drug*Time | Laposec | 2 | -0.1667 | |
| 22 | Drug*Time | Laposec | 2.5 | -0.1667 | |
| 23 | Drug*Time | Laposec | 3 | -0.1667 | |
| 24 | Drug*Time | Placebo | 0.5 | | -0.1667 |
| 25 | Drug*Time | Placebo | 1 | | -0.1667 |
| 26 | Drug*Time | Placebo | 1.5 | | -0.1667 |
| 27 | Drug*Time | Placebo | 2 | | -0.1667 |
| 28 | Drug*Time | Placebo | 2.5 | | -0.1667 |
| 29 | Drug*Time | Placebo | 3 | | -0.1667 |
| . | . | . | . | . | . |
| . | . | . | . | . | . |
| . | . | . | . | . | . |

(report continues with several pages of output)

The L matrices are used to form a linear combination of the betas corresponding to a specific hypothesis test or mean estimate. The L matrix in this example is used for testing whether there is a difference among the three levels of Drug.

### No.

This number is used for identifying the corresponding beta term.

### Effect

This column gives the model term.

### Factor Variables (e.g. Drug, Time)

These columns identify the level of each fixed effect to which the coefficients of the L matrix of the same line correspond.

### L1, L2, L3, …

L1, L2, L3, … are a group of column vectors that combine to form an L matrix. The L matrix in this example is used for testing whether there is a difference among the three levels of Drug.

## Discussion of Example 1 Results

The output shown for this example to this point has been for the full model with all interactions. It has been shown to illustrate the several sections of output that are available. In practice, when dealing with covariates, this model should be refined before making conclusions concerning the two drugs in question. The original F-test results are repeated below.

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Drug | 1.82 | 2 | 89.0 | 0.1677 |
| Time | 0.98 | 5 | 88.4 | 0.4358 |
| Cov | 3.30 | 1 | 87.1 | 0.0726 |
| Drug*Time | 0.86 | 10 | 87.0 | 0.5708 |
| Drug*Cov | 0.77 | 2 | 86.8 | 0.4662 |
| Time*Cov | 1.22 | 5 | 88.5 | 0.3078 |
| Drug*Time*Cov | 1.07 | 10 | 87.0 | 0.3947 |

These F-Values test Type-III (adjusted last) hypotheses.

Using a hierarchical step-down approach to model improvement, we begin by removing the highest order term, the three-way interaction (F-Value = 1.07, Prob Level = 0.3947). The F-test results for this new model are as follows.

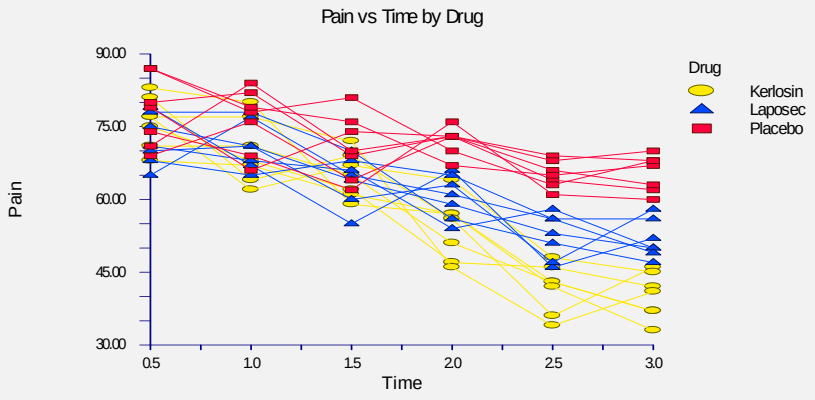
**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Drug | 4.17 | 2 | 96.8 | 0.0183 |
| Time | 1.34 | 5 | 98.0 | 0.2531 |
| Cov | 1.77 | 1 | 98.9 | 0.1866 |
| Drug*Time | 7.44 | 10 | 84.1 | 0.0000 |
| Drug*Cov | 2.23 | 2 | 92.4 | 0.1129 |
| Time*Cov | 2.37 | 5 | 98.0 | 0.0450 |

These F-Values test Type-III (adjusted last) hypotheses.

Since all interaction Prob Levels are now quite small, this model appears to be reasonable. Some researchers might argue to continue refinement by removing the Drug*Cov interaction (F-Value = 2.23, Prob Level = 0.1129). Such an argument is also reasonable, but this is not the course that is pursued here, since a moderately low prob level indicates there may be a mild Drug*Cov interaction effect.

The dominant prob level is the one associated with the Drug*Time interaction (F-Value = 7.44, Prob Level = 0.0000). This interaction can be clearly seen in the following scatter plot of the individual subjects. Note that the Placebo group does not decrease as rapidly as the Kerlosin group.




This interaction can be examined in greater detail by comparing the three levels of Drug at each time point (at the covariate value of 140).

**Individual Comparison Hypothesis Test Results**
**Covariates: Cov=140.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Time = 0.5, Drug: Kerlosin - Laposec | 6.16 | 4.33 | 1 | 100.0 | 0.0400 | 0.7206 [18] |
| Time = 0.5, Drug: Kerlosin - Placebo | -1.05 | 0.13 | 1 | 100.0 | 0.7205 | 1.0000 [18] |
| Time = 0.5, Drug: Laposec - Placebo | -7.21 | 6.37 | 1 | 100.0 | 0.0132 | 0.2370 [18] |
| Time = 1, Drug: Kerlosin - Laposec | 1.47 | 0.25 | 1 | 100.0 | 0.6161 | 1.0000 [18] |
| Time = 1, Drug: Kerlosin - Placebo | -7.29 | 6.21 | 1 | 99.9 | 0.0144 | 0.2583 [18] |
| Time = 1, Drug: Laposec - Placebo | -8.75 | 8.15 | 1 | 100.0 | 0.0052 | 0.0943 [18] |
| Time = 1.5, Drug: Kerlosin - Laposec | 2.35 | 0.72 | 1 | 99.9 | 0.3987 | 1.0000 [18] |
| Time = 1.5, Drug: Kerlosin - Placebo | -5.28 | 3.68 | 1 | 99.8 | 0.0578 | 1.0000 [18] |
| Time = 1.5, Drug: Laposec - Placebo | -7.63 | 7.57 | 1 | 99.9 | 0.0070 | 0.1267 [18] |
| Time = 2, Drug: Kerlosin - Laposec | -2.48 | 0.63 | 1 | 100.0 | 0.4277 | 1.0000 [18] |
| Time = 2, Drug: Kerlosin - Placebo | -14.12 | 19.44 | 1 | 100.0 | 0.0000 | 0.0005 [18] |
| Time = 2, Drug: Laposec - Placebo | -11.64 | 17.64 | 1 | 99.8 | 0.0001 | 0.0010 [18] |
| Time = 2.5, Drug: Kerlosin - Laposec | -11.05 | 16.57 | 1 | 99.7 | 0.0001 | 0.0017 [18] |
| Time = 2.5, Drug: Kerlosin - Placebo | -27.11 | 70.10 | 1 | 100.0 | 0.0000 | 0.0000 [18] |
| Time = 2.5, Drug: Laposec - Placebo | -16.06 | 26.37 | 1 | 100.0 | 0.0000 | 0.0000 [18] |
| Time = 3, Drug: Kerlosin - Laposec | -10.80 | 15.65 | 1 | 99.8 | 0.0001 | 0.0026 [18] |
| Time = 3, Drug: Kerlosin - Placebo | -25.19 | 84.92 | 1 | 99.8 | 0.0000 | 0.0000 [18] |
| Time = 3, Drug: Laposec - Placebo | -14.40 | 27.54 | 1 | 99.8 | 0.0000 | 0.0000 [18] |

The first Bonferroni-adjusted significant difference among levels of treatment occurs at Time = 2 hours. At Time = 2, the Kerlosin and Laposec means are significantly different from the Placebo mean (Bonferroni Prob Levels = 0.0005 and 0.0010, respectively), but not from each other (Bonferroni Prob Level = 1.0000). At times 2.5 hours and 3 hours all levels of Drug are significantly different, with Kerlosin showing the greatest pain reduction.

## Repeated and Random Component Specification

Another issue that should be considered from the beginning of the analysis is the covariance structure of the repeated measurements over time. The specification to this point involved both random (**G**) and the repeated (**R**) components of the model. The **G** and the **R** matrices are used to form the complete variance-covariance matrix of all the responses using the formula **V = ZGZ' + R**. The **G** and the **R** used to this point have the form

$$\mathbf{G} = \begin{pmatrix} \sigma_S^2 & & & \\ & \sigma_S^2 & & \\ & & \sigma_S^2 & \\ & & & \sigma_S^2 \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} \sigma^2 & & & & & \\ & \sigma^2 & & & & \\ & & \sigma^2 & & & \\ & & & \sigma^2 & & \\ & & & & \sigma^2 & \\ & & & & & \sigma^2 \end{pmatrix}$$

where G has dimension 21 by 21 and R has dimension 126 by 126. The resulting variance-covariance matrix, **V = ZGZ' + R**, has the form

**V** =

$$\begin{pmatrix} \sigma^2+\sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & 0 & 0 & \cdots \\ \sigma_S^2 & \sigma^2+\sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & 0 & 0 & \cdots \\ \sigma_S^2 & \sigma_S^2 & \sigma^2+\sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & 0 & 0 & \cdots \\ \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma^2+\sigma_S^2 & \sigma_S^2 & \sigma_S^2 & 0 & 0 & \cdots \\ \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma^2+\sigma_S^2 & \sigma_S^2 & 0 & 0 & \cdots \\ \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma_S^2 & \sigma^2+\sigma_S^2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2+\sigma_S^2 & \sigma_S^2 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_S^2 & \sigma^2+\sigma_S^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where each 6 by 6 block corresponds to a single patient. The full dimension of this matrix is 6*21 = 126 by 126.

The estimates of $\sigma_S^2$ and $\sigma^2$ for the model without the three-way interaction are 0.7063 and 24.6291, as shown in the output below.

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 0.7063 | Patient |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 24.6291 | Diagonal (Variance) |

The resulting 6 by 6 matrix for each subject (as shown in the output) is

**Estimated Vi Matrix of Subject = 1**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.3354 | 0.7063 | 0.7063 | 0.7063 | 0.7063 | 0.7063 |
| 2 | 0.7063 | 25.3354 | 0.7063 | 0.7063 | 0.7063 | 0.7063 |
| 3 | 0.7063 | 0.7063 | 25.3354 | 0.7063 | 0.7063 | 0.7063 |
| 4 | 0.7063 | 0.7063 | 0.7063 | 25.3354 | 0.7063 | 0.7063 |
| 5 | 0.7063 | 0.7063 | 0.7063 | 0.7063 | 25.3354 | 0.7063 |
| 6 | 0.7063 | 0.7063 | 0.7063 | 0.7063 | 0.7063 | 25.3354 |

The number 25.3354 comes from adding 0.7063 and 24.6291.

## Using Compound Symmetry as the Repeated Pattern Rather than Using a Random Component

An alternative specification that yields the same results is to remove the Random Component of the Model (Patient) and change the Repeated (Time) Covariance Pattern to Compound Symmetry. In this case, there is no **G** matrix and the **R** matrix has the form

$$\mathbf{R} = \begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

The true dimension of **R** is still 126 by 126 with 21 of the above matrices along the diagonal. The Repeated Component output becomes

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 25.3349 | Diagonal (Variance) |
| 1 | 1 | 0.0279 | Off-Diagonal (Correlation) |

Here, the estimate of $\sigma^2$ is 25.3349 and the estimate of $\rho$ is 0.0279.

The V matrix now has the form

$$\mathbf{V} = \begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & 0 & 0 & \cdots \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & 0 & 0 & \cdots \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & 0 & 0 & \cdots \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & 0 & 0 & \cdots \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & 0 & 0 & \cdots \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & \rho\sigma^2 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \rho\sigma^2 & \sigma^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and the estimated block for each subject using the compound symmetry specification is

**Estimated Vi Matrix of Subject = 1**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.3349 | 0.7067 | 0.7067 | 0.7067 | 0.7067 | 0.7067 |
| 2 | 0.7067 | 25.3349 | 0.7067 | 0.7067 | 0.7067 | 0.7067 |
| 3 | 0.7067 | 0.7067 | 25.3349 | 0.7067 | 0.7067 | 0.7067 |
| 4 | 0.7067 | 0.7067 | 0.7067 | 25.3349 | 0.7067 | 0.7067 |
| 5 | 0.7067 | 0.7067 | 0.7067 | 0.7067 | 25.3349 | 0.7067 |
| 6 | 0.7067 | 0.7067 | 0.7067 | 0.7067 | 0.7067 | 25.3349 |

which is identical (to rounding error) to the previous result using random and repeated component specification.

## Other Repeated Patterns (AR(1))

It is natural to expect that the covariances of measurements made closer together in time are more similar than those at more distant times. Several covariance pattern structures have been developed for such cases. A complete list of the available structures in the Mixed Models procedure is given elsewhere in the chapter. Here, we will examine one of the more common structures: AR(1).

Using the AR(1) covariance pattern, there are only two parameters, $\sigma^2$ and $\rho$, but the coefficient of $\sigma^2$ decreases exponentially as observations are farther apart. The **R** matrix has the form

$$\mathbf{R} = \begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 & \rho^4\sigma^2 & \rho^5\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 & \rho^4\sigma^2 \\ \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 & \rho^3\sigma^2 \\ \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 \\ \rho^4\sigma^2 & \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho^5\sigma^2 & \rho^4\sigma^2 & \rho^3\sigma^2 & \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

The true dimension of **R** is 126 by 126 with 21 of the above matrices along the diagonal.

The Repeated Component output becomes

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 25.3371 | Diagonal (Variance) |
| 1 | 1 | 0.0659 | Off-Diagonal (Correlation) |

Here, the estimate of $\sigma^2$ is 25.3371 and the estimate of $\rho$ is 0.0659.

The estimated block for each subject using the AR(1) specification is

**Estimated Vi Matrix of Subject = 1**

| Vi | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 25.3371 | 1.6692 | 0.1100 | 0.0072 | 0.0005 | 0.0000 |
| 2 | 1.6692 | 25.3371 | 1.6692 | 0.1100 | 0.0072 | 0.0005 |
| 3 | 0.1100 | 1.6692 | 25.3371 | 1.6692 | 0.1100 | 0.0072 |
| 4 | 0.0072 | 0.1100 | 1.6692 | 25.3371 | 1.6692 | 0.1100 |
| 5 | 0.0005 | 0.0072 | 0.1100 | 1.6692 | 25.3371 | 1.6692 |
| 6 | 0.0000 | 0.0005 | 0.0072 | 0.1100 | 1.6692 | 25.3371 |

The estimates of the covariance parameters using this formulation are closer to 0 as the time between measurements increases.

The AIC value may be used to compare the various covariance structures. The AIC value for the AR(1) specification is 725.77. The AIC value for the compound symmetry (and random component) specification is 725.94. A smaller AIC value indicates a better model. Thus, the AR(1) specification provides a slight improvement over the compound symmetry (and random component) specification.

## Example 2a – Two-Sample T-Test Assuming Equal Variance (One Between-Subject Factor, No Within-Subject Factors, No Covariates)

Examples 2a and 2b show how the use of the Mixed Models procedure gives the same results as the corresponding two-sided test in the Two-Sample T-Test procedure. Example 2c shows the extension to include a covariate in a two-sample test, which cannot be done using the Two-Sample T-Test procedure.

One of the simplest, yet very commonly used, designs is the two-group design. In this design, subjects are randomly assigned to, or randomly drawn from, one of two groups. A response is measured, and the means are compared. The common technique for analysis in this scenario is the two-sample (two-group) T-test. The data set-up for this design is two variables.

**TWOSAMPLE Dataset**

| Response | Treatment |
|---|---|
| 121 | Treatment |
| 105 | Treatment |
| 115 | Treatment |
| 130 | Treatment |
| 134 | Treatment |
| 136 | Treatment |
| 122 | Treatment |
| 114 | Treatment |
| . | . |
| . | . |
| . | . |
| 190 | Placebo |
| 186 | Placebo |
| 183 | Placebo |
| 175 | Placebo |

Using the T-Test – Two-Sample procedure, the two groups would be compared by entering *Response* as the Response Variable and *Treatment* as the Group Variable. An excerpt of the output appears as follows.

## Output Excerpt – Two-Sample T-Test Procedure

**Equal-Variance T-Test Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 |
|---|---|---|---|
| Difference <> 0 | 2.6278 | 0.012941 | Yes |

The equivalence of means is rejected (Prob Level = 0.012941) at the 0.05 alpha level.

The corresponding analysis in the Mixed Models procedure is similar, but an additional subject variable must be added. The subject variable identifies the subject to which each row belongs. When there are no repeated measurements, a subject variable may be created quickly by clicking any cell in a blank variable, selecting *Fill* from the *Edit* menu, and clicking on *Fill*.

**Two-Sample T-Test Example Dataset – TWOSAMPLE2**

| Response | Treatment | Subject |
|---|---|---|
| 121 | Treatment | 1 |
| 105 | Treatment | 2 |
| 115 | Treatment | 3 |
| 130 | Treatment | 4 |
| 134 | Treatment | 5 |
| 136 | Treatment | 6 |
| 122 | Treatment | 7 |
| 114 | Treatment | 8 |
| . | . | . |
| . | . | . |
| . | . | . |
| 190 | Placebo | 32 |
| 186 | Placebo | 33 |
| 183 | Placebo | 34 |
| 175 | Placebo | 35 |

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 2a** from the Template tab of the Mixed Models window.

### 1 Open the TWOSAMPLE2 dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **TWOSAMPLE2.s0**.
- Click **Open**.

### 2 Open the Mixed Models window.

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**. 'Response' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Subject** from the list of variables and then click **Ok**. 'Subject' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Treatment** for the **Factor (Categorical) Variables** text box.

**4 Specify the model.**

- Enter **Treatment** under **Model** for the **Fixed Effects Model**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Paramter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 288.5088 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Treatment | 6.91 | 1 | 33.0 | 0.0129 |

The Prob Level (0.0129) for the Treatment term is the same as the one given by the Two-Sample T-Test procedure. The estimate of the residual variance is 288.5088, which is the square of the standard deviation from the Two-Sample T-Test procedure.

# Example 2b – Two-Sample T-Test Assuming Unequal Variance (One Between-Subject Factor, No Within-Subject Factors, No Covariates)

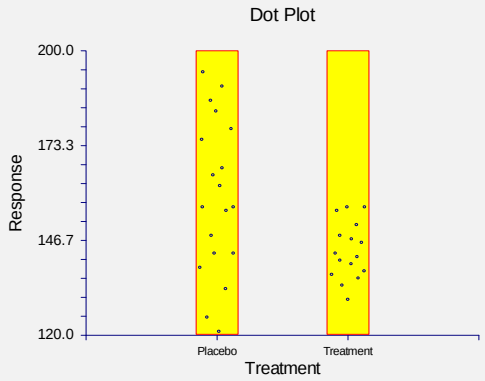
We now examine the TWOSAMPLE dataset without assuming equal variance among the two groups. The mean response is to be compared for a treatment and placebo.

**TWOSAMPLE Dataset**

| Response | Treatment |
|---|---|
| 121 | Treatment |
| 105 | Treatment |
| 115 | Treatment |
| 130 | Treatment |
| 134 | Treatment |
| 136 | Treatment |
| 122 | Treatment |
| 114 | Treatment |
| . | . |
| . | . |
| . | . |
| 190 | Placebo |
| 186 | Placebo |
| 183 | Placebo |
| 175 | Placebo |

The assumption of equal variance in this example is probably not a good one, as evidenced by the dot plot and equal variance tests shown below. The dot plot is obtained using the Dot Plots procedure, while the assumption tests are from the Two-Sample T-Test procedure.

**Dot Plot Section**




**Tests of Assumptions Section**

| Assumption | Value | Probability | Decision(.050) |
|---|---|---|---|
| Variance-Ratio Equal-Variance Test | 7.4226 | 0.000296 | Reject equal variances |
| Modified-Levene Equal-Variance Test | 11.8596 | 0.001579 | Reject equal variances |

The variance of the placebo group is much larger than that of the treatment group. The Aspin-Welch Unequal-Variance T-test should be used in place of the traditional T-test. An equivalent option is allowing for a separate variance for each group in the Mixed Models procedure.

Using the T-Test – Two-Sample procedure, the two groups would be compared by entering *Response* as the Response Variable and *Treatment* as the Group Variable. An excerpt of the output appears as follows.

## Unequal Variance Output Excerpt – Two-Sample T-Test Procedure

**Aspin-Welch Unequal-Variance Test Section**

| Alternative Hypothesis | T-Value | Prob Level | Reject H0 at .050 |
|---|---|---|---|
| Difference <> 0 | 2.8107 | 0.009801 | Yes |

The corresponding analysis in the Mixed Models procedure is similar, but an additional subject variable must be added. The subject variable identifies the subject to which each row belongs. When there are no repeated measurements, a subject variable may be created quickly by clicking any cell in a blank variable, selecting *Fill* from the *Edit* menu, and clicking on *Fill*.

**Two-Sample T-Test Example Dataset – TWOSAMPLE2**

| Response | Treatment | Subject |
|---|---|---|
| 121 | Treatment | 1 |
| 105 | Treatment | 2 |
| 115 | Treatment | 3 |
| 130 | Treatment | 4 |
| 134 | Treatment | 5 |
| 136 | Treatment | 6 |
| 122 | Treatment | 7 |
| 114 | Treatment | 8 |
| . | . | . |
| . | . | . |
| . | . | . |
| 190 | Placebo | 32 |
| 186 | Placebo | 33 |
| 183 | Placebo | 34 |
| 175 | Placebo | 35 |

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 2b** from the Template tab of the Mixed Models window. The difference in setup from Example 2a is the additional model specification in Step 4.

**1 Open the TWOSAMPLE2 dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **TWOSAMPLE2.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**. 'Response' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Subject** from the list of variables and then click **Ok**. 'Subject' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Treatment** for the **Factor (Categorical) Variables** text box.

**4 Specify the model.**

- Enter **Treatment** under **Model** for the **Fixed Effects Model**.
- Enter **Treatment** in **Groups** for the **Repeated (Time) Covariance Pattern**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Group Number | Estimated Value | Parameter Type |
|---|---|---|---|---|
| 1 | 1 | 1 | 475.5436 | Diagonal (Variance) |
| 1 | 1 | 2 | 64.0666 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Treatment | 7.90 | 1 | 23.5 | 0.0098 |

The Prob Level (0.0098) is the same as the one given by the Aspin-Welch Unequal Variance test of the Two-Sample T-Test procedure. The two variance estimates, 475.5436 and 64.0666, correspond to the placebo and treatment groups, respectively, and are the squares of the individual group standard deviations given in the Two-Sample T-Test procedure.

Examples 2a and 2b would likely be run using the Two-Sample T-Test procedure rather than the Mixed Models procedure. These examples are provided as an introduction to running the Mixed Models procedure for a simple case, as well as to show the flexibility of the Mixed Models procedure. Example 2c shows the extension of a two-sample test to the inclusion of a covariate. The Two-Sample T-Test procedure does not permit the inclusion of a covariate.

# Example 2c – Two-Sample T-Test with a Covariate (One Between-Subject Factor, No Within-Subject Factors, One Covariate)

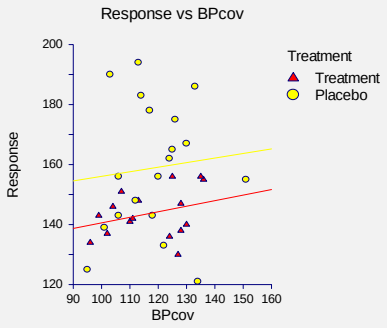
The two-sample analyses shown in Examples 2a and 2b would likely be carried out using the T-Test – Two-Sample procedure rather than the Mixed Models procedure because the T-Test procedure is easier to use and gives more specific output. However, when a covariate is measured for each subject there is no way to incorporate this into a simple T-test. The analysis becomes analysis of covariance, or ANCOVA. The General Linear Models or Multiple Regression procedures could be used, but in those, equal variances must be assumed. The flexibility we need for this analysis can only be achieved using the Mixed Models procedure. Adding a covariate only adds a couple of steps to the analysis without a covariate. The TWOSAMPLE2 dataset with the addition of a covariate, blood pressure (BP), becomes the TWOSAMPLECOV dataset.

**TWOSAMPLECOV Dataset**

| Response | Treatment | Subject | BPcov |
|---|---|---|---|
| 121 | Treatment | 1 | 110 |
| 105 | Treatment | 2 | 104 |
| 115 | Treatment | 3 | 128 |
| 130 | Treatment | 4 | 136 |
| 134 | Treatment | 5 | 96 |
| 136 | Treatment | 6 | 124 |
| 122 | Treatment | 7 | 111 |
| 114 | Treatment | 8 | 102 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| 190 | Placebo | 32 | 103 |
| 186 | Placebo | 33 | 133 |
| 183 | Placebo | 34 | 114 |
| 175 | Placebo | 35 | 126 |

A scatter plot of the two groups is obtained from the Scatter Plots procedure.

**Scatter Plot Section**



To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 2c** from the Template tab of the Mixed Models window.

**1 Open the TWOSAMPLECOV dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **TWOSAMPLECOV.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**. 'Response' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Subject** from the list of variables and then click **Ok**. 'Subject' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Treatment** for the **Factor (Categorical) Variables** text box.
- Select **BPcov** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Treatment BPcov** under **Model** for the **Fixed Effects Model**.
- Enter **Treatment** in **Groups** for the **Repeated (Time) Covariance Pattern**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output Excerpt

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Group Number | Estimated Value | Parameter Type |
|---|---|---|---|---|
| 1 | 1 | 1 | 475.0776 | Diagonal (Variance) |
| 1 | 1 | 2 | 61.4903 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Treatment | 7.71 | 1 | 23.2 | 0.0107 |
| BPcov | 1.55 | 1 | 18.7 | 0.2279 |

**Least Squares (Adjusted) Means**
**Covariates: BPcov=117.86**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 151.33 | 2.69 | 145.78 | 156.88 | 23.2 |
| **Treatment** | | | | | |
| Placebo | 158.79 | 5.00 | 148.29 | 169.30 | 18.0 |
| Treatment | 143.87 | 1.96 | 139.67 | 148.07 | 14.3 |

There is not strong evidence of a relationship between BPcov and Response (Prob Level = 0.2279). The difference in Treatment levels (Placebo and Treatment) is still seen (Prob Level = 0.0107). Because BPcov has little or no effect on the Response, the variance estimates for each group are similar to those obtain with including the covariate. Least squares adjusted means are given for the mean value of the covariate.

Comparisons and/or least square means could be obtained for any value of the covariate by specifying the desired value under the Covariates tab of the Mixed Models procedure. Specifying a covariate value of BPcov = 130 gives the following output.

**Individual Comparison Hypothesis Test Results**
**Covariates: BPcov=130.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Treatment | | 7.71 | 1 | 23.2 | 0.0107 | |
| Treatment: Placebo - Treatment | 14.92 | 7.71 | 1 | 23.2 | 0.0107 | 0.0107 [1] |

These F-Values test Type-III (adjusted last) hypotheses.

**Least Squares (Adjusted) Means**
**Covariates: BPcov=130.00**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 153.53 | 3.22 | 146.97 | 160.09 | 30.9 |
| **Treatment** | | | | | |
| Placebo | 160.99 | 5.28 | 150.02 | 171.95 | 21.1 |
| Treatment | 146.07 | 2.70 | 140.34 | 151.79 | 16.2 |

The F-test (F-Value = 7.71, Prob Level = 0.0107) is the same as the test for Treatment since the lines are assumed parallel in the fixed model. Thus, the F-test would be the same for any value of BPcov, unless the fixed model were changed to include the interaction Treatment*BPcov.

The least squares adjusted means (160.99 and 146.07) are adjusted to the covariate value of 130.

## Example 3a – One-Way ANOVA Design Assuming Equal Variance (One Between-Subject Factor, No Within-Subject Factors, No Covariates)

In a one-way layout design, two or more (usually three or more) groups are compared. Similar to the two-sample design, one column contains the response while another column identifies the groups. In this example, four plant food mixtures (salicylic acid, low iron, high iron, and no food) are compared in their ability to promote growth in beans. Twenty-eight plots are used in the experiment. The response is the weight of the beans harvested from the plot.

**BEAN Dataset**

| Food | Plot | Weight |
|---|---|---|
| Salicyl | 1 | 256 |
| Salicyl | 2 | 284 |
| Salicyl | 3 | 255 |
| Salicyl | 4 | 214 |
| Salicyl | 5 | 283 |
| Salicyl | 6 | 277 |
| Salicyl | 7 | 263 |
| LowIron | 8 | 293 |
| LowIron | 9 | 326 |
| LowIron | 10 | 313 |
| LowIron | 11 | 319 |
| LowIron | 12 | 321 |
| . | . | . |
| . | . | . |
| . | . | . |
| None | 26 | 238 |
| None | 27 | 259 |
| None | 28 | 243 |

This dataset could be analyzed using the One-Way Analysis of Variance procedure. The four groups would be compared by entering *Weight* as the Response Variable and *Food* as the Factor Variable.

## Output Excerpt – One-Way ANOVA Procedure

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level |
|---|---|---|---|---|---|
| A: Food | 3 | 64812.39 | 21604.13 | 20.68 | 0.000001* |
| S(A) | 24 | 25068.57 | 1044.524 | | |
| Total (Adjusted) | 27 | 89880.96 | | | |
| Total | 28 | | | | |

* Term significant at alpha = 0.05

The equivalence of means is rejected (F-Ratio = 20.68, Prob Level = 0.000001) at the 0.05 alpha level.

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 3a** from the Template tab of the Mixed Models window.

**1 Open the BEAN dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **BEAN.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Weight** from the list of variables and then click **Ok**. 'Weight' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Plot** from the list of variables and then click **Ok**. 'Plot' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Food** for the **Factor (Categorical) Variables** text box.

**4 Specify the model.**

- Enter **Food** under **Model** for the **Fixed Effects Model**.

**5 Specify the comparisons.**

- On the Mixed Models window, select the **Comparisons tab**.
- Select **All Pairs** under **Comparison** for **Default Factor Comparisons**.

**6 Specify the reports.**

- Leave all reports and plots at their default values.

**7 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

# Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 1044.5237 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Food | 20.68 | 3 | 24.0 | 0.0000 |

These F-Values test Type-III (adjusted last) hypotheses.

**Individual Comparison Hypothesis Test Results**

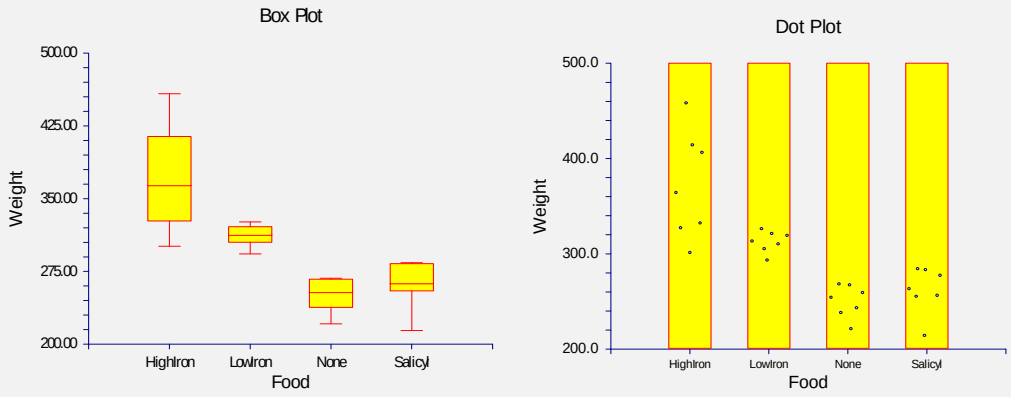
| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Food | | 20.68 | 3 | 24.0 | 0.0000 | |
| Food: HighIron - LowIron | 59.29 | 11.78 | 1 | 24.0 | 0.0022 | 0.0131 [6] |
| Food: HighIron - None | 121.71 | 49.64 | 1 | 24.0 | 0.0000 | 0.0000 [6] |
| Food: HighIron - Salicyl | 110.00 | 40.54 | 1 | 24.0 | 0.0000 | 0.0000 [6] |
| Food: LowIron - None | 62.43 | 13.06 | 1 | 24.0 | 0.0014 | 0.0083 [6] |
| Food: LowIron - Salicyl | 50.71 | 8.62 | 1 | 24.0 | 0.0072 | 0.0434 [6] |
| Food: None - Salicyl | -11.71 | 0.46 | 1 | 24.0 | 0.5042 | 1.0000 [6] |

The overall F-test (F-Value = 20.68, Prob Level = 0.0000) comparing the means indicates there is strong evidence for differences among means. The individual comparison tests, with appropriate Bonferroni adjustments for multiple testing, indicate there are differences in means among all levels except between None and Salicylic acid.

The overall F-test (F-Value = 20.68, Prob Level = 0.0000) is identical to the one that results from the One-Way ANOVA procedure. If equal variances can reasonably assumed, the One-Way ANOVA procedure gives more detailed information than this procedure and should be used instead. However, for the case of unequal variances among groups, there is no way to use the One-Way Analysis of Variance procedure to analyze this dataset, except possibly with a transformation. In this example, the issue of unequal variances is important since the test for equal variance is rejected (see below).

The following output is generated from the One-Way Analysis of Variance and Dot Plots procedures.

**Plots**



**Tests of Assumptions Section**

| Assumption | Test Value | Prob Level | Decision (0.05) |
|---|---|---|---|
| Modified-Levene Equal-Variance Test | 6.2834 | 0.002667 | Reject |

It appears HighIron group has a much larger variance than the other groups.

Comparing the levels of Food assuming different variance within each group using the Mixed Models procedure is shown in the next example.

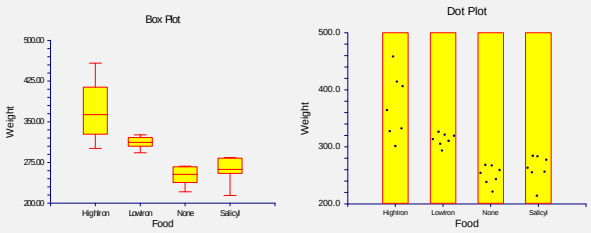
# Example 3b – One-Way ANOVA Design Assuming Unequal Variance (One Between-Subject Factor, No Within-Subject Factors, No Covariates)

In this example, four plant food mixtures (salicylic acid, low iron, high iron, and no food) are compared in their ability to promote growth in beans. Twenty-eight plots are used in the experiment. The response is the weight of the beans harvested from the plot. This example differs from the previous example in that the assumption of equal variances between groups is removed.

**BEAN Dataset**

| Food | Plot | Weight |
|---|---|---|
| Salicyl | 1 | 256 |
| Salicyl | 2 | 284 |
| Salicyl | 3 | 255 |
| Salicyl | 4 | 214 |
| Salicyl | 5 | 283 |
| Salicyl | 6 | 277 |
| Salicyl | 7 | 263 |
| LowIron | 8 | 293 |
| LowIron | 9 | 326 |
| LowIron | 10 | 313 |
| LowIron | 11 | 319 |
| LowIron | 12 | 321 |
| . | . | . |
| . | . | . |
| . | . | . |
| None | 26 | 238 |
| None | 27 | 259 |
| None | 28 | 243 |

In this example, there is no way to use the One-Way Analysis of Variance procedure to analyze this dataset if unequal variances are assumed. The issue of unequal variances is important since an equal variance test is rejected (See the plots and equal variance test of the BEAN data below). The following output is generated from the One-Way Analysis of Variance and Dot Plots procedures.




**Tests of Assumptions Section**

| Assumption | Test Value | Prob Level | Decision (0.05) |
|---|---|---|---|
| Modified-Levene Equal-Variance Test | 6.2834 | 0.002667 | Reject |

It appears HighIron group has a much larger variance than the other groups.

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 3b** from the Template tab of the Mixed Models window. The only difference in specification from the previous example is the additional entry in Step 4.

### 1 Open the BEAN dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **BEAN.s0**.
- Click **Open**.

### 2 Open the Mixed Models window.

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Weight** from the list of variables and then click **Ok**. 'Weight' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Plot** from the list of variables and then click **Ok**. 'Plot' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Food** for the **Factor (Categorical) Variables** text box.

### 4 Specify the model.

- Enter **Food** under **Model** for the **Fixed Effects Model**.
- Enter **Food** in **Groups** for the **Repeated (Time) Covariance Pattern**.

**5 Specify the comparisons.**

- On the Mixed Models window, select the **Comparisons tab**.
- Select **All Pairs** under **Comparison** for **Default Factor Comparisons**.

**6 Specify the reports.**

- Leave all reports and plots at their default values.

**7 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Group Number | Estimated Value | Parameter Type |
|---|---|---|---|---|
| 1 | 1 | 1 | 3174.2381 | Diagonal (Variance) |
| 1 | 1 | 2 | 123.2857 | Diagonal (Variance) |
| 1 | 1 | 3 | 290.6667 | Diagonal (Variance) |
| 1 | 1 | 4 | 589.9048 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Food | 27.13 | 3 | 10.8 | 0.0000 |

**Individual Comparison Hypothesis Test Results**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Food | | 27.13 | 3 | 10.8 | 0.0000 | |
| Food: HighIron - LowIron | 59.29 | 7.46 | 1 | 6.5 | 0.0317 | 0.1899 [6] |
| Food: HighIron - None | 121.71 | 29.93 | 1 | 7.1 | 0.0009 | 0.0054 [6] |
| Food: HighIron - Salicyl | 110.00 | 22.50 | 1 | 8.2 | 0.0014 | 0.0083 [6] |
| Food: LowIron - None | 62.43 | 65.90 | 1 | 10.3 | 0.0000 | 0.0001 [6] |
| Food: LowIron - Salicyl | 50.71 | 25.24 | 1 | 8.4 | 0.0009 | 0.0053 [6] |
| Food: None - Salicyl | -11.71 | 1.09 | 1 | 10.8 | 0.3192 | 1.0000 [6] |

The overall F-test comparing the means indicates there is strong evidence for differences among means. The individual comparison tests indicate there are differences in means among all levels except between None and Salicylic acid.

The Repeated Component Parameter Estimates section shows a different residual variance estimate for each of the four groups. The first variance estimate (3174.2381), corresponding to the high iron food mixture, is much larger than the others.

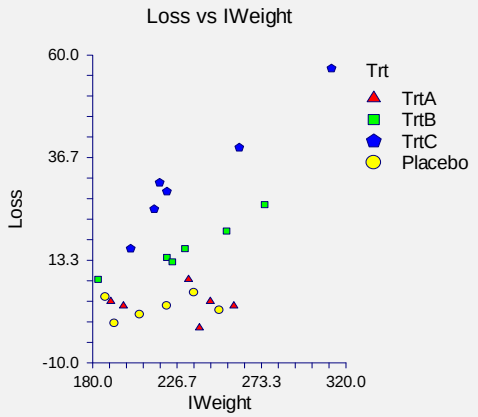
# Example 4 – ANCOVA Design (One Between-Subject Factor, No Within-Subject Factors, One Covariate)

In this example, three weight loss treatments (and a placebo) are compared. Twenty-four patients are randomly assigned to the three treatments and the placebo. Weight loss is the response. The weight of each participant before treatment is measured as a covariate. The researchers wish to compare the levels of Treatment at low (190 lbs.), medium (230 lbs.), and high (270 lbs.) values of initial weight.

**WEIGHTLOSS Dataset**

| Trt | Patient | IWeight | Loss |
|---|---|---|---|
| TrtA | 1 | 197 | 3 |
| TrtA | 2 | 245 | 4 |
| TrtA | 3 | 233 | 9 |
| TrtA | 4 | 239 | -2 |
| TrtA | 5 | 258 | 3 |
| TrtA | 6 | 190 | 4 |
| TrtB | 7 | 221 | 14 |
| TrtB | 8 | 231 | 16 |
| TrtB | 9 | 224 | 13 |
| TrtB | 10 | 183 | 9 |
| TrtB | 11 | 275 | 26 |
| TrtB | 12 | 254 | 20 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| Placebo | 19 | 187 | 5 |
| Placebo | 20 | 192 | -1 |
| Placebo | 21 | 250 | 2 |
| Placebo | 22 | 236 | 6 |
| Placebo | 23 | 221 | 3 |
| Placebo | 24 | 206 | 1 |

A scatter plot of the data is shown below.



This analysis could be run using the Multiple Regression procedure by entering Loss as the Dependent Variable, IWeight as a Numeric Independent Variable and Trt as a Categorical Independent Variable. The Default Contrast Type is set to Standard Set. The Custom Model is Trt|Iweight.

### Output Excerpt – Multiple Regression Procedure

**Analysis of Variance Detail Section**

| Model Term | DF | R2 | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (5%) |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | | 4592.667 | 4592.667 | | | |
| Model | 7 | 0.9682 | 4695.169 | 670.7384 | 69.613 | 0.0000 | 1.0000 |
| IWeight | 1 | 0.0631 | 306.1491 | 306.1491 | 31.774 | 0.0000 | 0.9995 |
| Trt | 3 | 0.0331 | 160.7248 | 53.57495 | 5.560 | 0.0083 | 0.8710 |
| IWeight*Trt | 3 | 0.0822 | 398.6192 | 132.8731 | 13.790 | 0.0001 | 0.9990 |
| Error | 16 | 0.0318 | 154.1646 | 9.635287 | | | |
| Total(Adjusted) | 23 | 1.0000 | 4849.333 | 210.8406 | | | |

The significant IWeight*Trt interaction (F-Ratio = 13.790, Prob Level = 0.0001) indicates there are differences among the slopes of the treatment groups. These results will be compared to those of the Mixed Models procedure in the output and discussion that follows.

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 4** from the Template tab of the Mixed Models window.

**1 Open the WEIGHTLOSS dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **WEIGHTLOSS.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Loss** from the list of variables and then click **Ok**. 'Loss' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Patient** from the list of variables and then click **Ok**. 'Patient' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Trt** for the **Factor (Categorical) Variables** text box.
- Select **IWeight** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Trt IWeight Trt*IWeight** under **Model** for the **Fixed Effects Model**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 9.6353 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Trt | 5.56 | 3 | 16.0 | 0.0083 |
| IWeight | 31.77 | 1 | 16.0 | 0.0000 |
| Trt*IWeight | 13.79 | 3 | 16.0 | 0.0001 |

**Individual Comparison Hypothesis Test Results**
**Covariates: IWeight=227.83**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Trt | | 89.78 | 3 | 16.0 | 0.0000 | |
| Trt: Placebo - TrtA | -0.49 | 0.06 | 1 | 16.0 | 0.8022 | 1.0000 [3] |
| Trt: Placebo - TrtB | -12.69 | 43.24 | 1 | 16.0 | 0.0000 | 0.0000 [3] |
| Trt: Placebo - TrtC | -26.61 | 186.07 | 1 | 16.0 | 0.0000 | 0.0000 [3] |

These F-Values test Type-III (adjusted last) hypotheses.

**Least Squares (Adjusted) Means**
**Covariates: IWeight=227.83**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 12.95 | 0.66 | 11.54 | 14.35 | 16.0 |
| **Trt** | | | | | |
| Placebo | 3.00 | 1.45 | -0.07 | 6.07 | 16.0 |
| TrtA | 3.49 | 1.27 | 0.80 | 6.18 | 16.0 |
| TrtB | 15.69 | 1.28 | 12.98 | 18.39 | 16.0 |
| TrtC | 29.61 | 1.31 | 26.84 | 32.39 | 16.0 |

The Term-by-Term Hypothesis Test Results are identical to those given in the Multiple Regression procedure output.

The Prob Level for the interaction Trt*IWeight confirms what is seen in the scatter plot: the slopes differ for the different treatments. Two important sections of the output that are available in the Mixed Models procedure that are not available in the Multiple Regression procedure are mean comparisons and least squares (adjusted) means at specific values of the covariates.

The Individual Comparison Hypothesis Tests of the preceding output, however, are not very useful. They compare the placebo to each of the treatments at the mean of the covariate (IWeight = 227.83). To better understand the nature of the interaction, it is useful to compare the placebo to the three treatments at various values of the covariate. Some caution should be exercised with the number of values that are chosen because the probability levels (P-values) are adjusted according to the number of comparisons that are tested. Generally a low, medium, and high value of the covariate should suffice.

The researchers wished to compare the levels of Treatment at low (190 lbs.), medium (230 lbs.), and high (270 lbs.) values of initial weight. Such values should be chosen prior to collecting the data or at least before looking at the results. Comparisons and means at these values of the covariate are output by selecting the Covariate tab, selecting IWeight under Covariate Variable, and entering 190 230 270 under Compute Means at these Values. The relevant output is shown in the section that follows.

**Individual Comparison Hypothesis Test Results**
**Covariates: IWeight=190.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Trt | | 11.30 | 3 | 16.0 | 0.0003 | 0.0009 [3] |
| Trt: Placebo - TrtA | -1.94 | 0.43 | 1 | 16.0 | 0.5219 | 1.0000 [9] |
| Trt: Placebo - TrtB | -6.71 | 5.26 | 1 | 16.0 | 0.0357 | 0.3215 [9] |
| Trt: Placebo - TrtC | -15.22 | 29.96 | 1 | 16.0 | 0.0001 | 0.0005 [9] |

These F-Values test Type-III (adjusted last) hypotheses.

**Individual Comparison Hypothesis Test Results**
**Covariates: IWeight=230.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Trt | | 94.27 | 3 | 16.0 | 0.0000 | 0.0000 [3] |
| Trt: Placebo - TrtA | -0.41 | 0.04 | 1 | 16.0 | 0.8394 | 1.0000 [9] |
| Trt: Placebo - TrtB | -13.03 | 43.69 | 1 | 16.0 | 0.0000 | 0.0001 [9] |
| Trt: Placebo - TrtC | -27.26 | 188.32 | 1 | 16.0 | 0.0000 | 0.0000 [9] |

These F-Values test Type-III (adjusted last) hypotheses.

**Individual Comparison Hypothesis Test Results**
**Covariates: IWeight=270.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Trt | | 79.33 | 3 | 16.0 | 0.0000 | 0.0000 [3] |
| Trt: Placebo - TrtA | 1.12 | 0.07 | 1 | 16.0 | 0.7907 | 1.0000 [9] |
| Trt: Placebo - TrtB | -19.35 | 24.24 | 1 | 16.0 | 0.0002 | 0.0014 [9] |
| Trt: Placebo - TrtC | -39.30 | 112.97 | 1 | 16.0 | 0.0000 | 0.0000 [9] |

These F-Values test Type-III (adjusted last) hypotheses.

**Least Squares (Adjusted) Means**
**Covariates: IWeight=190.00**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 7.96 | 1.06 | 5.72 | 10.19 | 16.0 |
| **Trt** | | | | | |
| Placebo | 1.99 | 1.90 | -2.03 | 6.01 | 16.0 |
| TrtA | 3.93 | 2.27 | -0.88 | 8.73 | 16.0 |
| TrtB | 8.70 | 2.23 | 3.98 | 13.42 | 16.0 |
| TrtC | 17.21 | 2.03 | 12.90 | 21.52 | 16.0 |

**Least Squares (Adjusted) Means**
**Covariates: IWeight=230.00**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 13.23 | 0.67 | 11.81 | 14.65 | 16.0 |
| **Trt** | | | | | |
| Placebo | 3.06 | 1.51 | -0.14 | 6.26 | 16.0 |
| TrtA | 3.47 | 1.28 | 0.76 | 6.17 | 16.0 |
| TrtB | 16.09 | 1.27 | 13.40 | 18.78 | 16.0 |
| TrtC | 30.32 | 1.29 | 27.58 | 33.06 | 16.0 |

**Least Squares (Adjusted) Means**
**Covariates: IWeight=270.00**

| Name | Mean | Standard Error of Mean | 95.0% Lower Conf. Limit for Mean | 95.0% Upper Conf. Limit for Mean | DF |
|---|---|---|---|---|---|
| **Intercept** | | | | | |
| Intercept | 18.51 | 1.24 | 15.88 | 21.14 | 16.0 |
| **Trt** | | | | | |
| Placebo | 4.13 | 3.30 | -2.87 | 11.13 | 16.0 |
| TrtA | 3.00 | 2.53 | -2.35 | 8.36 | 16.0 |
| TrtB | 23.48 | 2.13 | 18.96 | 27.99 | 16.0 |
| TrtC | 43.43 | 1.66 | 39.90 | 46.96 | 16.0 |

Examination of the individual comparison hypothesis tests shows that the mean difference from the placebo for those with a higher initial weight is greater than the mean difference for those with a lower initial weight, with the exception of Treatment A, for which there is no significant improvement in weight loss over the placebo.

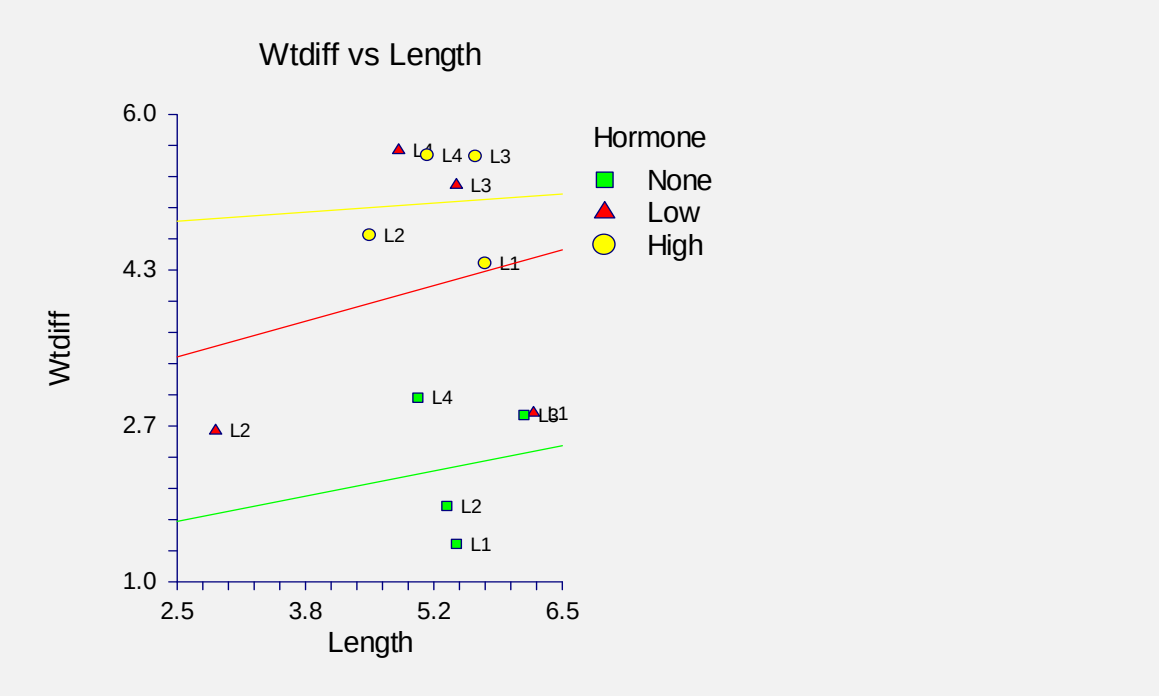
# Example 5 – Factorial Design (Two Between-Subject Factors, No Within-Subject Factors, One Covariate)

In a factorial design, more than one fixed factor is analyzed in a single experiment. One variable contains the response and two or more other variables identify the groups. In this example, a study is conducted to determine the effect of a growth hormone on trout growth at fish hatcheries. Twelve fish are compared in the study. Each of the 12 fish receives a different combination of hormone dose (none, low, or high) and amount of fish food (Level 1, Level 2, Level 3, or Level 4). The response is increase in weight after 3 weeks in the tank. The length of each fish prior to treatment is measured as a covariate.

**FISH Dataset**

| Fish | Food | Hormone | Length | Wtdiff |
|---|---|---|---|---|
| 1 | Level1 | None | 5.4 | 1.408 |
| 2 | Level1 | Low | 6.2 | 2.808 |
| 3 | Level1 | High | 5.7 | 4.407 |
| 4 | Level2 | None | 5.3 | 1.813 |
| 5 | Level2 | Low | 2.9 | 2.618 |
| 6 | Level2 | High | 4.5 | 4.708 |
| 7 | Level3 | None | 6.1 | 2.786 |
| 8 | Level3 | Low | 5.4 | 5.247 |
| 9 | Level3 | High | 5.6 | 5.551 |
| 10 | Level4 | None | 5.0 | 2.971 |
| 11 | Level4 | Low | 4.8 | 5.618 |
| 12 | Level4 | High | 5.1 | 5.563 |

A scatter plot of the data is shown below.



This analysis could be run using the Multiple Regression procedure by entering Wtdiff as the Dependent Variable, Length as a Numeric Independent Variable, and Food and Hormone as Categorical Independent Variables. The Default Contrast Type is set to Standard Set. The Custom Model is Food+Hormone+Length+Hormone*Length.

### Output Excerpt – Multiple Regression Procedure

**Analysis of Variance Detail Section**

| Model Term | DF | R2 | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (5%) |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | | 172.5057 | 172.5057 | | | |
| Model | 8 | 0.9426 | 25.00806 | 3.126008 | 6.155 | 0.0813 | 0.5013 |
| Food | 3 | 0.3088 | 8.192474 | 2.730825 | 5.377 | 0.1003 | 0.4209 |
| Hormone | 2 | 0.0042 | 0.1111072 | 5.555359E-02 | 0.109 | 0.8998 | 0.0571 |
| Length | 1 | 0.0002 | 5.378415E-03 | 5.378415E-03 | 0.011 | 0.9245 | 0.0506 |
| Hormone*Length | 2 | 0.0028 | 7.335056E-02 | 3.667528E-02 | 0.072 | 0.9319 | 0.0547 |
| Error | 3 | 0.0574 | 1.523545 | 0.5078484 | | | |
| Total(Adjusted) | 11 | 1.0000 | 26.53161 | 2.411964 | | | |

None of the terms of the model are significant in this example. However, the model can be refined. These results will be compared to those of the Mixed Models procedure in the output and discussion that follows.

To run the analysis in the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 5** from the Template tab of the Mixed Models window.

**1 Open the FISH dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **FISH.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Wtdiff** from the list of variables and then click **Ok**. 'Wtdiff' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Fish** from the list of variables and then click **Ok**. 'Fish' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Food** and **Hormone** for the **Factor (Categorical) Variables** text box.
- Select **Length** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Food Hormone Length Hormone*Length** under **Model** for the **Fixed Effects Model**.

**5 Specify the comparisons.**

- On the Mixed Models window, select the **Comparisons tab**.
- Select **All Pairs** under **Comparison** for **Default Factor Comparisons**.

**6 Specify the reports.**

- Leave all reports and plots at their default values.

**7 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 0.5078 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Food | 5.38 | 3 | 3.0 | 0.1003 |
| Hormone | 0.11 | 2 | 3.0 | 0.8998 |
| Length | 0.01 | 1 | 3.0 | 0.9245 |
| Hormone*Length | 0.07 | 2 | 3.0 | 0.9319 |

These F-Values test Type-III (adjusted last) hypotheses.

None of the factors are significant in this example. The model should be refined before examining the remainder of the output. It is evident that the length of the fish prior to treatment has little effect on the response, or at least with this small sample size the effect of length is not detectable. Removing the two non-signficant terms associated with the length covariate and re-running the analysis gives the output that follows.

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 0.3076 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Food | 9.09 | 3 | 6.0 | 0.0119 |
| Hormone | 26.49 | 2 | 6.0 | 0.0011 |

These F-Values test Type-III (adjusted last) hypotheses.

**Individual Comparison Hypothesis Test Results**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Food | | 9.09 | 3 | 6.0 | 0.0119 | |
| Food: Level1 - Level2 | -0.17 | 0.14 | 1 | 6.0 | 0.7172 | 1.0000 [6] |
| Food: Level1 - Level3 | -1.65 | 13.33 | 1 | 6.0 | 0.0107 | 0.0641 [6] |
| Food: Level1 - Level4 | -1.84 | 16.56 | 1 | 6.0 | 0.0066 | 0.0395 [6] |
| Food: Level2 - Level3 | -1.48 | 10.71 | 1 | 6.0 | 0.0170 | 0.1020 [6] |
| Food: Level2 - Level4 | -1.67 | 13.62 | 1 | 6.0 | 0.0102 | 0.0613 [6] |
| Food: Level3 - Level4 | -0.19 | 0.17 | 1 | 6.0 | 0.6904 | 1.0000 [6] |
| | | | | | | |
| Hormone | | 26.49 | 2 | 6.0 | 0.0011 | |
| Hormone: High - Low | 0.98 | 6.30 | 1 | 6.0 | 0.0459 | 0.1376 [3] |
| Hormone: High - None | 2.81 | 51.44 | 1 | 6.0 | 0.0004 | 0.0011 [3] |
| Hormone: Low - None | 1.83 | 21.73 | 1 | 6.0 | 0.0035 | 0.0104 [3] |

With the removal of the covariate terms, the results now show strong evidence of differences between levels of Food (F-Value = 9.09, Prob Level = 0.0119) and Hormone dose (F-Value = 26.49, Prob Level = 0.0011). Individual comparisons indicate evidence that the Level 1 mean is different from the Level 4 mean (Bonferroni Prob Level = 0.0395). The High and the Low levels of Hormone are significantly different (Bonferroni Prob Levels = 0.0011 and 0.0104, respectively) from the level None.

Similar results can be obtained using the General Linear Models procedure. However, the General Linear Models procedure would not allow the user to model different variances among groups, if this were desired.

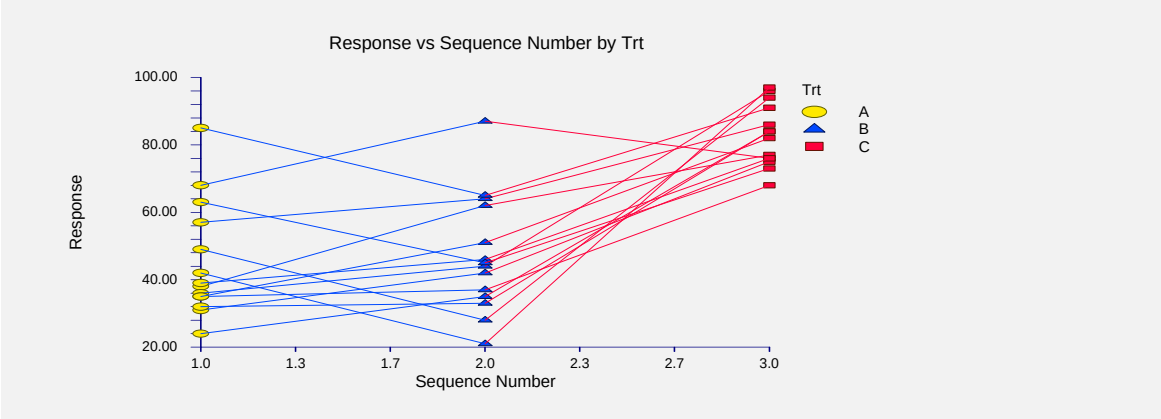
# Example 6 – Randomized Complete Block Design (No Between-Subject Factors, One Within-Subject Factor, No Covariates)

In a study to compare 3 treatments, three patients from each of 14 doctors are randomly assigned to each of the three treatments. A single response is measured for each patient following treatment. The result is a randomized complete block design with 14 blocks (doctors). The goal is to determine whether there are any differences among the three treatments. The data are contained in the RCBD data set.

**RCBD Dataset**

| Doctor | Patient | Trt | Response |
|---|---|---|---|
| 1 | 1 | A | 57 |
| 1 | 2 | B | 64 |
| 1 | 3 | C | 86 |
| 2 | 4 | A | 85 |
| 2 | 5 | B | 65 |
| 2 | 6 | C | 91 |
| 3 | 7 | A | 24 |
| 3 | 8 | B | 35 |
| 3 | 9 | C | 84 |
| 4 | 10 | A | 68 |
| 4 | 11 | B | 87 |
| 4 | 12 | C | 76 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| 13 | 37 | A | 49 |
| 13 | 38 | B | 28 |
| 13 | 39 | C | 94 |
| 14 | 40 | A | 32 |
| 14 | 41 | B | 33 |
| 14 | 42 | C | 84 |

A plot showing the 3 patients for each doctor is shown below.




Because there are no covariates, this analysis could be run using the Repeated Measures Analysis of Variance procedure by entering Response as the Response Variable, Doctor as the Subject

Variable, and Trt as Within Factor 1. The Model Specification is Full model except subject interactions combined with error.

## Output Excerpt – Repeated Measures Analysis of Variance Procedure

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A: Doctor | 13 | 4232.119 | 325.5476 | 1.83 | 0.091201 | |
| B: Trt | 2 | 12507.19 | 6253.595 | 35.23 | 0.000000* | 1.000000 |
| S | 26 | 4614.81 | 177.4927 | | | |
| Total (Adjusted) | 41 | 21354.12 | | | | |
| Total | 42 | | | | | |

* Term significant at alpha = 0.05

The difference in Treatment levels is highly significant (F-Ratio = 35.23, Prob Level = 0.000000). These results will be compared to those of the Mixed Models procedure in the output and discussion that follows.

To run the analysis using the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 6** from the Template tab of the Mixed Models window.

### 1 Open the RCBD dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **RCBD.s0**.
- Click **Open**.

### 2 Open the Mixed Models window.

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Response** from the list of variables and then click **Ok**. 'Response' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Doctor** from the list of variables and then click **Ok**. 'Doctor' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Trt** for the **Factor (Categorical) Variables** text box.

### 4 Specify the model.

- Enter **Trt** under **Model** for the **Fixed Effects Model**.
- Enter **Doctor** under **Model** for the **Random Model (Subject Terms Only)**.

### 5 Specify the comparisons.

- On the Mixed Models window, select the **Comparisons tab**.
- Select **All Pairs** under **Comparison** for **Default Factor Comparisons**.

### 6 Specify the reports.

- Leave all reports and plots at their default values.

### 7 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 49.3517 | Doctor |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 177.4927 | Diagonal (Variance) |

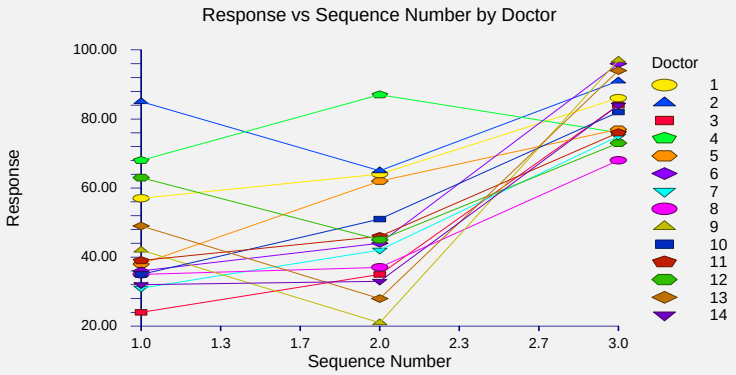
**Term-by-Term Hypothesis Test Results**

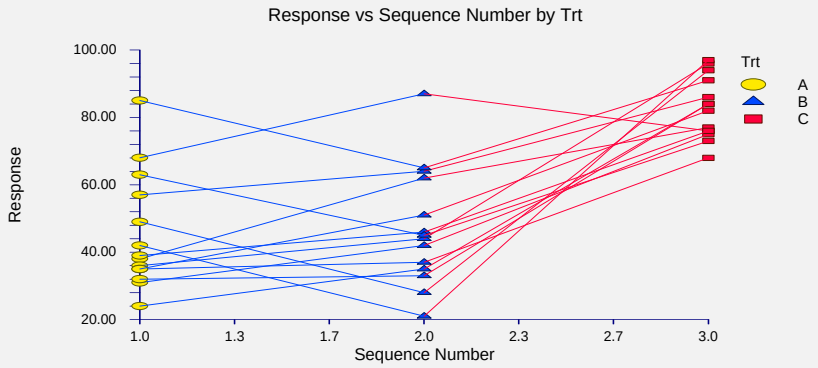
| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Trt | 35.23 | 2 | 26.0 | 0.0000 |

**Individual Comparison Hypothesis Test Results**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Trt | | 35.23 | 2 | 26.0 | 0.0000 | |
| Trt: A - B | -1.86 | 0.14 | 1 | 26.0 | 0.7153 | 1.0000 [3] |
| Trt: A - C | -37.50 | 55.46 | 1 | 26.0 | 0.0000 | 0.0000 [3] |
| Trt: B - C | -35.64 | 50.10 | 1 | 26.0 | 0.0000 | 0.0000 [3] |

**Subject Plots**






The results of this test match those of the Repeated Measures ANOVA procedure (F-Value = 35.23, Prob Level = 0.0000). All the reports indicate that the mean response for Treatment C is much higher than A and B (the Bonferroni Prob Levels for A vs. C and B vs. C are both extremely small).

The second subject plot seems to indicate that the variation within responses of Treatment C is considerably smaller than the variation within Treatments A and B. This can be accounted for by entering Trt for Groups in the Repeated Variance Pattern on the Variables tab.

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 8.1829 | Doctor |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Group Number | Estimated Value | Parameter Type |
|---|---|---|---|---|
| 1 | 1 | 1 | 273.4668 | Diagonal (Variance) |
| 1 | 1 | 2 | 298.3705 | Diagonal (Variance) |
| 1 | 1 | 3 | 77.7942 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Trt | 41.30 | 2 | 17.2 | 0.0000 |

The conclusions do not change when the unequal variance is accounted for, but the estimated variances are indeed quite different across treatments. The estimated variances for Treatments A and B are 273.4668 and 298.3705, respectively, while the estimated variance for Treatment C is only 77.7942. These tests based on unequal variance assumptions are more accurate than those where equal variances were assumed.

# Example 7 – Complex Split-Plot Design (One Between-Subject Factor, Two Within-Subject Factors, Two Covariates)

In a standard split-plot design, plots are randomized to a between-plot treatment and are also sub-divided, with each sub-division receiving a different within-plot treatment. This example involves a more complex split-design with an additional within-plot factor and two covariates.

In a study to compare the effectiveness of 3 tutoring methods, 84 students (42 male, 42 female) are randomly assigned to 14 tutors (7 graduates, 7 undergraduates) in groups of 6 (3 male, 3 female). Each tutor uses a different tutoring method for each student according to the scheme below. A pre-exam is administered to each student before the semester of tutoring begins. IQ is also obtained for each student. The response is the score on an exam taken at the end of the semester.

**TUTOR Dataset**

| Educ | Tutor | Student | Method | Gender | Preexam | IQ | Exam |
|---|---|---|---|---|---|---|---|
| Undergr | 1 | 1 | A | M | 45 | 117 | 70 |
| Undergr | 1 | 2 | A | F | 35 | 113 | 84 |
| Undergr | 1 | 3 | B | M | 68 | 94 | 95 |
| Undergr | 1 | 4 | B | F | 47 | 103 | 77 |
| Undergr | 1 | 5 | C | M | 25 | 100 | 78 |
| Undergr | 1 | 6 | C | F | 24 | 95 | 88 |
| Undergr | 2 | 7 | A | M | 16 | 96 | 80 |
| Undergr | 2 | 8 | A | F | 38 | 99 | 75 |
| Undergr | 2 | 9 | B | M | 59 | 98 | 77 |
| Undergr | 2 | 10 | B | F | 75 | 105 | 82 |
| Undergr | 2 | 11 | C | M | 65 | 106 | 76 |
| Undergr | 2 | 12 | C | F | 45 | 98 | 94 |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| Grad | 14 | 79 | A | M | 27 | 109 | 77 |
| Grad | 14 | 80 | A | F | 36 | 104 | 81 |
| Grad | 14 | 81 | B | M | 24 | 79 | 84 |
| Grad | 14 | 82 | B | F | 27 | 99 | 72 |
| Grad | 14 | 83 | C | M | 33 | 93 | 75 |
| Grad | 14 | 84 | C | F | 39 | 109 | 63 |

The only procedure that can be used to incorporate all the variables of this analysis in a single model is the Mixed Models procedure.

To run the analysis using the Mixed Models procedure, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 7** from the Template tab of the Mixed Models window.

**1 Open the TUTOR dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **TUTOR.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Exam** from the list of variables and then click **Ok**. 'Exam' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Tutor** from the list of variables and then click **Ok**. 'Tutor' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Educ**, **Method**, and **Gender** for the **Factor (Categorical) Variables** text box.
- Select **Preexam** and **IQ** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Educ Method Gender Preexam IQ** under **Model** for the **Fixed Effects Model**.
- Enter **Tutor** under **Model** for the **Random Model (Subject Terms Only)**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

# Mixed Models Output

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 0.8116 | Tutor |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 79.4848 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Educ | 0.68 | 1 | 14.3 | 0.4241 |
| Method | 1.00 | 2 | 65.8 | 0.3747 |
| Gender | 0.21 | 1 | 65.4 | 0.6455 |
| Preexam | 0.05 | 1 | 76.1 | 0.8194 |
| IQ | 1.07 | 1 | 77.0 | 0.3049 |

There is no statistical evidence of differences among the levels of Education, Method, or Gender (all Prob Levels > 0.05).

# Example 8 – Cross-Over Design (No Between-Subject Factors, Two Within-Subject Factors, One Covariate)

In a basic two-level cross-over design, each subject receives both treatments, but (approximately) half receive the two treatments in the opposite order. In this example, researchers are comparing two drugs for their effect on heart rate in rats. Each rat is given both drugs, with a short washout period between drug administrations, but the order of the drugs is reversed in half of the rats. An initial heart rate (IHR) measurement is taken immediately before administration of each of the drugs.

**CROSS Dataset**

| Rat | Period | Trtcross | IHR | HR |
|---|---|---|---|---|
| 1 | 1 | Drug A | 389 | 357 |
| 1 | 2 | Drug B | 383 | 381 |
| 2 | 1 | Drug B | 372 | 409 |
| 2 | 2 | Drug A | 390 | 385 |
| 3 | 1 | Drug A | 396 | 386 |
| 3 | 2 | Drug B | 372 | 377 |
| 4 | 1 | Drug B | 389 | 376 |
| 4 | 2 | Drug A | 398 | 385 |
| 5 | 1 | Drug A | 404 | 396 |
| 5 | 2 | Drug B | 378 | 370 |
| 6 | 1 | Drug B | 394 | 394 |
| 6 | 2 | Drug A | 392 | 366 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 18 | 1 | Drug B | 382 | 381 |
| 18 | 2 | Drug A | 396 | 380 |
| 19 | 1 | Drug A | 380 | 391 |
| 19 | 2 | Drug B | 387 | 392 |
| 20 | 1 | Drug B | 408 | 403 |
| 20 | 2 | Drug A | 391 | 371 |

To run the analysis, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 8** from the Template tab of the Mixed Models window.

### 1 Open the CROSS dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **CROSS.s0**.
- Click **Open**.

### 2 Open the Mixed Models window.

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **HR** from the list of variables and then click **Ok**. 'HR' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Rat** from the list of variables and then click **Ok**. 'Rat' will appear in the Subject Variable box.
- Make sure there is no entry in the **Time Variable** box.
- Select **Period** and **Trtcross** for the **Factor (Categorical) Variables** text box.
- Select **IHR** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Trtcross Period IHR** under **Model** for the **Fixed Effects Model**.
- Enter **Rat** under **Model** for the **Random Model (Subject Terms Only)**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixed Models Output

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 6.9397 | Rat |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 189.6138 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Trtcross | 3.99 | 1 | 20.6 | 0.0592 |
| Period | 0.49 | 1 | 17.5 | 0.4932 |
| IHR | 2.00 | 1 | 35.6 | 0.1659 |

The F-test for Trtcross is nearly significant (F-value = 3.99, Prob Level = 0.0592) at the 0.05 level. There appears to be no period effect (F-value = 0.49, Prob Level = 0.4932) nor relationship between the initial heart rate (F-value = 3.99, Prob Level = 0.0592) and the response heart rate.

The advantages of using mixed models in cross-over designs are usually more pronounced when there is missing data. Missing values often occur in cross-over designs when subjects fail to appear for the second treatment. Another advantage of mixed models in cross-over designs over conventional analyses occurs when there are three or more treatments involved. In such cases, the

cross-over design may be considered a repeated measures design, and specific covariate patterns can be used to model the similarity in repeated measurements. That is, measurements that are taken closer together may be expected to vary more similarly, while measurements at distant periods may not. The Mixed Models procedure provides greater flexibility in modeling options for such situations.

## Example 9 – Random Coefficients Model (One Between-Subject Factor, No Within-Subject Factors, One Covariate, Unequal Time Points, Missing Data)

Researchers would like to determine the effect of a new hair loss treatment. Eighteen men are randomly divided into two groups. One group receives the placebo (shampoo without treatment), the other group receives the hair loss treatment (shampoo with treatment). The participants are asked to shampoo daily and return to the lab after every two months for one year. At each visit, participants are given a hair re-growth score. As is sometimes the case with human subjects, the return visits were not as scheduled. Some participants returned before or after the scheduled two month period, while some others dropped out of the study.

**HAIR Dataset**

| Treatment | Individual | Time | Regrowth |
|---|---|---|---|
| Placebo | 1 | 2 | 14 |
| Placebo | 1 | 4 | 5 |
| Placebo | 1 | 7 | 3 |
| Placebo | 1 | 9 | 7 |
| Placebo | 1 | 12 | 8 |
| Placebo | 2 | 2 | 6 |
| Placebo | 2 | 4 | 3 |
| Placebo | 3 | 2 | 5 |
| Placebo | 3 | 5 | 2 |
| Placebo | 3 | 10 | 4 |
| Placebo | 4 | 2 | 7 |
| Placebo | 4 | 4 | 9 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| Trt | 17 | 6 | 7 |
| Trt | 17 | 12 | 3 |
| Trt | 18 | 2 | 7 |
| Trt | 18 | 4 | 10 |
| Trt | 18 | 6 | 14 |
| Trt | 18 | 8 | 17 |

To run the analysis, you may enter the values according to the instructions below (beginning with Step 3) or load the completed template **Example 9** from the Template tab of the Mixed Models window.

**1 Open the HAIR dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **HAIR.s0**.
- Click **Open**.

**2 Open the Mixed Models window.**

- On the menus, select **Analysis**, then **Mixed Models**. The Mixed Models procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Mixed Models window, select the **Variables tab**.
- Double-click in the **Response Variable** text box. This will bring up the variable selection window.
- Select **Regrowth** from the list of variables and then click **Ok**. 'Regrowth' will appear in the Response Variable box.
- Double-click in the **Subject Variable** text box. This will bring up the variable selection window.
- Select **Individual** from the list of variables and then click **Ok**. 'Individual' will appear in the Subject Variable box.
- Select **Time** for the **Time Variable** box.
- Select **Treatment** for the **Factor (Categorical) Variables** text box.
- Select **Time** for the **Covariate (Continuous) Variables** text box.

**4 Specify the model.**

- Enter **Treatment Time Treatment*Time** under **Model** for the **Fixed Effects Model**.
- Enter **Individual Individual*Time** under **Model** for the **Random Model (Subject Terms Only)**.
- Check the box next to **Covariances** under **Random Model (Subject Terms Only)**.

**5 Specify the reports.**

- Leave all reports and plots at their default values.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

# Mixed Models Output

**Random Component Parameter Estimates (G Matrix)**

| Component Number | Parameter Number | Estimated Value | Model Term |
|---|---|---|---|
| 1 | 1 | 16.6318 | Individual |
| 1 | 2 | 1.8015 | Individual*Time |
| 1 | 3 | -4.2599 | Individual, Individual*Time |

**Repeated Component Parameter Estimates (R Matrix)**

| Component Number | Parameter Number | Estimated Value | Parameter Type |
|---|---|---|---|
| 1 | 1 | 5.9013 | Diagonal (Variance) |

**Term-by-Term Hypothesis Test Results**

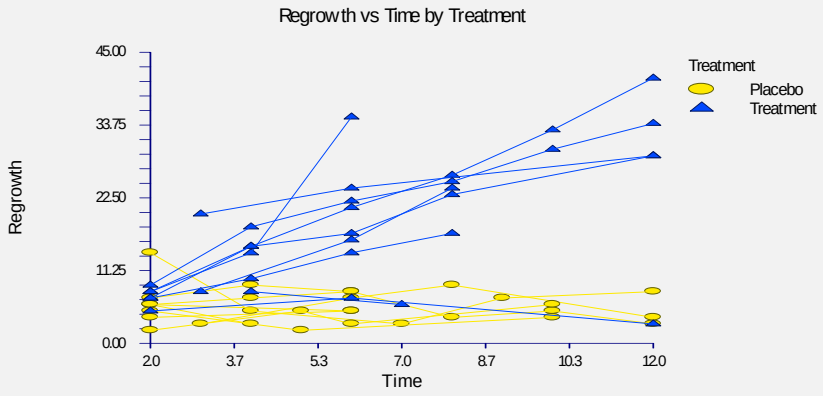
| Model Term | F-Value | Num DF | Denom DF | Prob Level |
|---|---|---|---|---|
| Treatment | 0.20 | 1 | 13.6 | 0.6585 |
| Time | 7.74 | 1 | 12.3 | 0.0162 |
| Treatment*Time | 9.50 | 1 | 12.3 | 0.0093 |

These F-Values test Type-III (adjusted last) hypotheses.

**Individual Comparison Hypothesis Test Results**
**Covariates: Time=6.00**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| Treatment | | 18.72 | 1 | 15.1 | 0.0006 | |
| Treatment: Placebo - Trt | -11.84 | 18.72 | 1 | 15.1 | 0.0006 | 0.0006 [1] |

These F-Values test Type-III (adjusted last) hypotheses.




The significant Treatment*Time interaction (F-Value = 9.50, Prob Level = 0.0093) indicates that the differences between the Treatment and the Placebo are different at different times. If comparisons are made at times 2, 7, and 12, the results are

**Individual Comparison Hypothesis Test Results**

| Comparison/ Covariate(s) | Comparison Mean Difference | F-Value | Num DF | Denom DF | Raw Prob Level | Bonferroni Prob Level |
|---|---|---|---|---|---|---|
| **Treatment: Placebo - Treatment** | | | | | | |
| Time=2.00 | -3.21 | 3.93 | 1 | 14.4 | 0.0669 | 0.2007 [3] |
| Time=7.00 | -14.00 | 17.56 | 1 | 14.5 | 0.0008 | 0.0025 [3] |
| Time=12.00 | -24.80 | 13.89 | 1 | 13.2 | 0.0025 | 0.0074 [3] |

There is strong evidence that the difference in means increases as time increases.

## Chapter 230

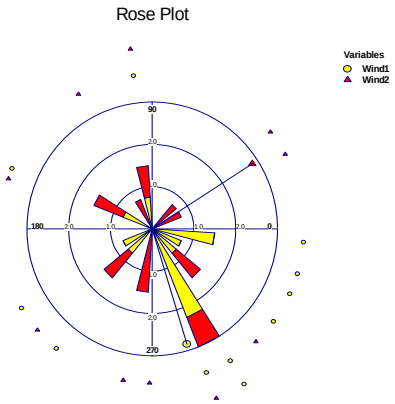
# Circular Data Analysis

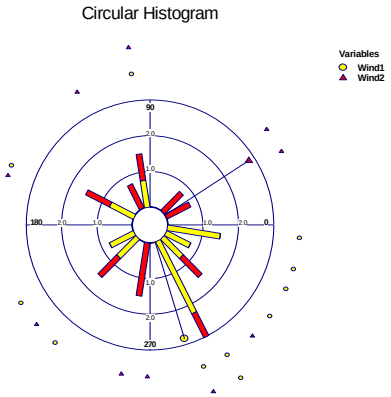
## Introduction

This procedure computes summary statistics, generates rose plots and circular histograms, computes hypothesis tests appropriate for one, two, and several groups, and computes the circular correlation coefficient for circular data.

Angular data, recorded in degrees or radians, is generated in a wide variety of scientific research areas. Examples of angular (and cyclical) data include daily wind directions, ocean current directions, departure directions of animals, direction of bone-fracture plane, and orientation of bees in a beehive after stimuli.

The usual summary statistics, such as the sample mean and standard deviation, cannot be used with angular values. For example, consider the average of the angular values 1 and 359. The simple average is 180. But with a little thought, we might conclude that 0 is a better answer. Because of this and other problems, a special set of techniques have been developed for analyzing angular data. This procedure implements many of those techniques.




## Technical Details

Suppose a sample of *n* angles $a_1, a_2, ..., a_n$ is to be summarized. It is assumed that these angles are in degrees. Fisher (1993) and Mardia & Jupp (2000) contain definitions of various summary statistics that are used for angular data. These results will be presented next. Let

$$C_p = \sum_{i=1}^{n} \cos(pa_i),\ \overline{C}_p = \frac{C_p}{n},\ S_p = \sum_{i=1}^{n} \sin(pa_i),\ \overline{S}_p = \frac{S_p}{n},$$

$$R_p = \sqrt{C_p^2 + S_p^2},\ \overline{R}_p = \frac{R_p}{n}$$

$$T_p = \begin{cases} \tan^{-1}\left(\dfrac{\overline{S}_p}{\overline{C}_p}\right) & \overline{C}_p > 0, \overline{S}_p > 0 \\ \tan^{-1}\left(\dfrac{\overline{S}_p}{\overline{C}_p}\right) + \pi & \overline{C}_p < 0 \\ \tan^{-1}\left(\dfrac{\overline{S}_p}{\overline{C}_p}\right) + 2\pi & \overline{S}_p < 0, \overline{C}_p > 0 \end{cases}$$

To interpret these quantities it may be useful to imagine that each angle represents a vector of length one in the direction of the angle. Suppose these individual vectors are arranged so that the beginning of the first vector is at the origin, the beginning of the second vector is at the end of the first, the beginning of the third vector is at the end of the second, and so on. We can then imagine a single vector $\vec{a}$ that will stretch from the origin to the end of the last observation.

$R_1$, called the *resultant length*, is the length of $\vec{a}$. $\overline{R}_1$ is the *mean resultant length* of $\vec{a}$. Note that $\overline{R}_1$ varies between zero and one and that a value of $\overline{R}_1$ near one implies that there was little variation in values of the angles.

The *mean direction*, $\theta$, is a measure of the mean of the individual angles. $\theta$ is estimated by $T_1$.

The *circular variance*, *V*, measures the variation in the angles about the mean direction. *V* varies from zero to one. The formula for *V* is

$$V = 1 - \overline{R}_1$$

The *circular standard deviation*, *v*, is defined as

$$v = \sqrt{-2\ln\left(\overline{R}_1\right)}$$

The *circular dispersion*, used in the calculation of confidence intervals, is defined as

$$\delta = \frac{1 - T_2}{2\overline{R}_1^2}$$

The *skewness* is defined as

$$s = \frac{\overline{R}_2 \sin(T_2 - 2T_1)}{\left(1 - \overline{R}_1\right)^{3/2}}$$

The *kurtosis* is defined as

$$k = \frac{\overline{R}_2 \cos(T_2 - 2T_1) - \overline{R}_1^4}{\left(1 - \overline{R}_1\right)^2}$$

## Correction for Grouped Data

When the angles are grouped, a multiplicative correction for *R* may be necessary. The corrected value is given by

$$\overline{R}_p^* = g\overline{R}_p$$

where

$$g = \frac{\pi / J}{\sin(\pi / J)}$$

Here *J* is the number of equi-sized arcs. Thus, for monthly data, *J* would be 12.

## Confidence Interval for the Mean Direction

Upton & Fingleton (1989) page 220 give a confidence interval for the mean direction when no distributional assumption is made as

$$T_1 \pm \sin^{-1}\left(z_{\alpha/2}\hat{\sigma}\right)$$

where

$$\hat{\sigma} = \sqrt{\frac{n(1-H)}{4R^2}}$$

$$H = \frac{1}{n}\left\{\cos(2T_1)\sum_{i=1}^{n}\cos(2a_i) + \sin(2T_1)\sum_{i=1}^{n}\sin(2a_i)\right\}$$

## Circular Uniform Distribution

*Uniformity* refers to the situation in which all values around the circle are equally likely. The probability distribution on a circle with this property is the *circular uniform distribution,* or simply, the uniform distribution. The probability density function is given by

$$f(a) = \frac{1}{360}$$

The probability between any two points is given by

$$\Pr(a_1 < a_2 | a_1 \le a_2, a_2 \le a_1 + 2\pi) = \frac{a_2 - a_1}{360}$$

### Tests of Uniformity

*Uniformity* refers to the situation in which all values around the circle are equally likely. Occasionally, it is useful to perform a statistical test of whether a set of data do not follow the uniform distribution. Several tests of uniformity have been developed. Note that when any of the following tests are rejected, we can conclude that the data were not uniform. However, when the test is not rejected, we cannot conclude that the data follow the uniform distribution. Rather, we do not have enough evidence to reject the null hypothesis of uniformity.

## Rayleigh Test

The Rayleigh test, discussed in Mardia & Jupp (2000) pages 94-95, is the score test and the likelihood ratio test for uniformity within the von Mises distribution family. The Rayleigh test statistic is $2n\overline{R}^2$. For large samples, the distribution of this statistic under uniformity is a chi-square with two degrees of freedom with an error of approximation of $O\left(n^{-1}\right)$. A closer approximation to the chi-square with two degrees of freedom is achieved by the modified Rayleigh test. This test, which has an error of $O\left(n^{-2}\right)$, is calculated as follows.

$$S^* = \left(1 - \frac{1}{2n}\right)2n\overline{R}^2 + \frac{n\overline{R}^4}{2}$$

## Modified Kuiper's Test

The modified Kuiper's test, Mardia & Jupp (2000) pages 99-103, was designed to test uniformity against any alternative. It measures the distance between the cumulative uniform distribution function and the empirical distribution function. It is accurate for samples as small as 8. The test statistic, *V*, is calculated as follows

$$V = V_n\left(\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}}\right)$$

where

$$V_n = \max_{i=1 \text{ to } n}\left(\frac{a_{(i)}}{360} - \frac{i}{n}\right) - \min_{i=1 \text{ to } n}\left(\frac{a_{(i)}}{360} - \frac{i}{n}\right) + \frac{1}{n}$$

Published critical values of *V* are

| ***V*** | **Alpha** |
|---|---|
| 1.537 | 0.150 |
| 1.620 | 0.100 |
| 1.747 | 0.050 |
| 1.862 | 0.025 |
| 2.001 | 0.010 |

This table was used to create an interpolation formula from which the alpha values are calculated.

## Watson Test

The following uniformity test is outlined in Mardia & Jupp pages 103-105. The test is conducted by calculating $U^2$ and comparing it to a table of values. If the calculated value is greater than the critical value, the null hypothesis of uniformity is rejected. Note that the test is only valid for samples of at least eight angles.

The calculation of $U^2$ is as follows

$$U^2 = \sum_{i=1}^{n}\left[u_{(i)} - \frac{i - \frac{1}{2}}{n} - \overline{u} + \frac{1}{2}\right]^2 + \frac{1}{12n}$$

where

$$\bar{u} = \frac{\sum_{i=1}^{n} u_{(i)}}{n}, \; u_{(i)} = \frac{a_{(i)}}{360}$$

$a_{(1)} \le a_{(2)} \le a_{(3)} \le \cdots \le a_{(n)}$ are the sorted angles. Note that maximum likelihood estimates of $\kappa$ and $\theta$ are used in the distribution function. Mardia & Jupp (2000) present a table of critical values that has been entered into ***NCSS***. When a value of $U^2$ is calculated, the table is interpolated to determine its significance level.

Published critical values of $U^2$ are

| $U^2$ | Alpha |
|---|---|
| 0.131 | 0.150 |
| 0.152 | 0.100 |
| 0.187 | 0.050 |
| 0.221 | 0.025 |
| 0.267 | 0.010 |

## Von Mises Distributions

The *Von Mises distribution* takes the role in circular statistics that is held by the normal distribution in standard linear statistics. In fact, it is shaped like the normal distribution, except that its tails are truncated.

The probability density function is given by

$$f(a;\theta,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\left[\kappa \cos(a-\theta)\right]$$

where $I_p(x)$ (the modified Bessel function of the first kind and order p) is defined by

$$I_p(x) = \sum_{r=0}^{\infty} \frac{1}{(r+p)!r!}\left(\frac{x}{2}\right)^{2r+p}, \quad p = 0,1,2,\ldots$$

In particular

$$I_0(x) = \sum_{r=0}^{\infty} \frac{1}{(r!)^2}\left(\frac{x}{2}\right)^{2r}$$
$$= \frac{1}{2\pi}\int_0^{2\pi} e^{x\cos(\theta)}d\theta$$

The parameter $\theta$ is the *mean direction* and the parameter $\kappa$ is the *concentration parameter*.

The distribution is unimodal. It is symmetric about *A*. It appears as a normal distribution that is truncated at plus and minus 180 degrees. When $\kappa$ is zero, the von Mises distribution reduces to the uniform distribution. As $\kappa$ gets large, the von Mises distribution approaches the normal distribution.

## Point Estimation

The maximum likelihood estimate of $\theta$ is the sample mean direction. That is, $\hat{\theta} = T_1$.

The maximum likelihood of $\kappa$ is the solution to

$$A_1(\kappa) = \overline{R}$$

where

$$A_1(x) = \frac{I_1(x)}{I_0(x)}.$$

That is, the MLE of $\kappa$ is given by

$$\kappa^* = A_1^{-1}\left(\overline{R}\right)$$

This can be approximated by (see Fisher (1993) page 88 and Mardia & Jupp (2000) pages 85-86)

$$\kappa^* = \begin{cases} 2\overline{R} + \overline{R}^3 + \dfrac{5\overline{R}^5}{6} & \overline{R} < 0.53 \\ -0.4 + 1.39\overline{R} + \dfrac{0.43}{1-\overline{R}} & 0.53 \le \overline{R} < 0.53 \\ \dfrac{1}{3\overline{R} - 4\overline{R}^2 + \overline{R}^3} & \overline{R} \ge 0.85 \end{cases}$$

This estimate is very biased. This bias is corrected by using the following modified estimator.

$$\hat{\kappa} = \begin{cases} \begin{cases} \max\left(\kappa^* - \dfrac{2}{n\kappa^*}, 0\right) & \kappa^* < 2 \\ \dfrac{(n-1)^3\kappa^*}{n\left(n^2+1\right)} & \kappa^* \ge 2 \end{cases} & n \le 15 \\ \kappa^* & n > 15 \end{cases}$$

## Test for a Specified Mean Direction of Von Mises Data

There are several different hypothesis tests that have been proposed for testing $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, where $\theta_0$ is a specific value of the mean direction. The tests presented here require the additional assumption that the data follow the Von Mises distribution, at least approximately.

It will be useful to adopt the following notation.

$$\overline{C}^* = \frac{1}{n}\sum_{i=1}^{n}\cos\left(a_i - \theta_0\right)$$

$$\overline{S}^* = \frac{1}{n}\sum_{i=1}^{n}\sin\left(a_i - \theta_0\right)$$

$$\overline{R}^* = \sqrt{\left[\overline{S}^*\right]^2 + \left[\overline{C}^*\right]^2}$$

### Score Test

The score test, given by Mardia & Jupp (2000) page 123, is computed as

$$\chi_S^2 = \frac{n\hat{\kappa}}{A_1(\hat{\kappa})}\left(\overline{S}^*\right)^2$$

For large *n*, $\chi_S^2$ follows the chi-square distribution with one degree of freedom.

### Likelihood Ratio Test

The likelihood ratio test, given by Mardia & Jupp (2000) page 122, is computed as

$$\chi_L^2 = \begin{cases} \dfrac{4n\left[\left(\overline{R}^*\right)^2 - \left(\overline{C}^*\right)^2\right]}{2 - \left(\overline{C}^*\right)^2} & \text{if } n \geq 5 \text{ and } \overline{C}^* \leq 2/3 \\ \dfrac{2n^3}{n^2 + \left(n\overline{C}^*\right)^2 + 3n} \log\left(\dfrac{1 - \left(\overline{C}^*\right)^2}{1 - \left(\overline{R}^*\right)^2}\right) & \text{if } n \geq 5 \text{ and } \overline{C}^* > 2/3 \end{cases}$$

The test statistic, $\chi_L^2$, follows a chi-square distribution with one degree of freedom.

### Watson & Williams Test

The Watson and Williams test, given by Mardia & Jupp (2000) page 123, is computed as

$$F = \frac{\overline{R}^* - \overline{C}^*}{\left(1 - \overline{R}^*\right)/(n-1)} \text{ if } \overline{C}^* \geq 5/6$$

The test statistic, *F*, follows an *F* distribution with one and *n*-1 degrees of freedom.

### Stephens Test

This test, given by Fisher (1993) pages 93-94, is computed as

$$E = \frac{\sin\left(T_1 - \theta_0\right)}{\sqrt{1/\left(n\hat{\kappa}\overline{R}\right)}}$$

If $\hat{\kappa} \geq 2$, *E* follows the standard normal distribution.

## Confidence Interval for Mean Direction assuming Von Mises

A general confidence interval for $\theta$ was given above. When the data can be assumed to follow a von Mises distribution, a more appropriate interval is given by Mardia & Jupp (2000) page 124 and Upton & Fingleton (1989) page 269. This confidence interval is given by

$$T_1 \pm \cos^{-1}\left(\sqrt{\frac{2n\left[2R^2 - nz_\alpha^2\right]}{R^2\left(4n - z_\alpha^2\right)}}\right) \quad \text{if } \overline{R} \le 2/3$$

$$T_1 \pm \cos^{-1}\left(\frac{\sqrt{n^2 - \left(n^2 - R^2\right)\exp\left(\frac{z_\alpha^2}{n}\right)}}{R}\right) \quad \text{if } \overline{R} > 2/3$$

## Test for a Specified Concentration of Von Mises Data

Suppose you want to test a one-sided hypothesis concerning $\kappa$, given that the data come from a Von Mises distribution and that the mean direction parameter is unknown. Fisher (1993) page 95 suggests the following procedure when $\hat{\kappa} \ge 2$.

When testing $\kappa = \kappa_0$ versus $\kappa < \kappa_0$, reject the null hypothesis if

$$\overline{R} < 1 - \frac{\chi^2_{n-1;\alpha}}{2n}\left(\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}\right)$$

When testing $\kappa = \kappa_0$ versus $\kappa > \kappa_0$, reject the null hypothesis if

$$\overline{R} > 1 - \frac{\chi^2_{n-1;1-\alpha}}{2n}\left(\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}\right)$$

These tests are based on the result that

$$\frac{2n\left(1-\overline{R}\right)}{\frac{1}{\kappa_0} + \frac{3}{8\kappa_0^2}} \sim \chi^2_{n-1}$$

## Confidence Interval for Concentration of Von Mises

An approximate confidence interval for $\kappa$ when $\hat{\kappa} > 2$ was given by Mardia & Jupp (2000) pages 126-127 as

$$\left(\frac{1+\sqrt{1+3b}}{4b}, \frac{1+\sqrt{1+3d}}{4d}\right)$$

where

$$b = \frac{n\left(1-\overline{R}\right)}{\chi^2_{n-1,1-\alpha/2}}$$

$$d = \frac{n\left(1-\overline{R}\right)}{\chi^2_{n-1,\alpha/2}}$$

## Goodness of Fit Tests for the Von Mises Distribution

### Stephens Test

The following goodness-of-fit test, published by Lockhart & Stephens (1985) as a modification of the Watson test for the circle, is outlined in Fisher (1993) page 84. The test is conducted by calculating $U^2$ and comparing it to a table of values. If the calculated value is greater than the critical value, the null hypothesis of Von Misesness is rejected. Note that the test is only valid for samples of at least 20 angles.

The calculation of $U^2$ is as follows

$$U^2 = \sum_{i=1}^{n}\left[\hat{p}_{(i)} - \frac{2i-1}{2n}\right]^2 - n\left(\bar{\hat{p}} - \frac{1}{2}\right)^2 + \frac{1}{12n}$$

where

$$\bar{\hat{p}} = \frac{\sum_{i=1}^{n}\hat{p}_{(i)}}{n}$$

$$\hat{p}_{(i)} = F_{\kappa}\left(a_{(i)} - T_1\right)$$

$a_{(1)} \le a_{(2)} \le a_{(3)} \le \cdots \le a_{(n)}$ are the sorted angles and $F_{\kappa}\left(a-\theta\right)$ is the cumulative distribution function of the von Mises distribution. Note that maximum likelihood estimates of $\kappa$ and $\theta$ are used in the distribution function. Lockhart & Stephens (1985) present a table of critical values that has been entered into ***NCSS***. When a value of $U^2$ is calculated, the table is interpolated to determine its significance level.

### Cox Test

Mardia & Jupp (2000) pages 142-143 present a von Mises goodness-of-fit test that was originally given by Cox (1975).

The test statistic, *C*, is distributed as a chi-squared variable with two degrees of freedom under the null hypothesis that the data follow the von Mises distribution. It is calculated as follows.

$$C = \frac{s_c^2}{n v_c\left(\hat{\kappa}\right)} + \frac{s_s^2}{n v_s\left(\hat{\kappa}\right)}$$

where

$$s_c = \sum_{i=1}^{n}\cos 2\left(a_i - T_1\right) - n\alpha_2\left(\hat{\kappa}\right)$$

$$s_s = \sum_{i=1}^{n}\sin 2\left(a_i - T_1\right)$$

$$v_c\left(x\right) = \frac{1+\alpha_4}{2} - \alpha_2^2 - \frac{\left[\alpha_1/2 + \alpha_3/2 - \alpha_1\alpha_2\right]^2}{\left(1+\alpha_2\right)/2 - \alpha_1^2}$$

$$v_S(x) = \frac{\alpha_1 - \alpha_4}{2} - \frac{(\alpha_1 - \alpha_3)^2}{1 - \alpha_2}$$

## Multi-Group Tests

Three multi-group tests are available for testing hypotheses about two or more groups. The nonparametric uniform-scores test tests whether the distributions of the groups are identical. The Watson-Williams F test tests whether a set of mean directions are equal given that the concentrations are unknown, but equal, given that the groups each follow a von Mises distribution. The concentration homogeneity test tests whether the concentration parameters are equal, given that the groups each follow a von Mises distribution.

### Mardia-Watson-Wheeler Uniform-Scores Test

Suppose you have *g* populations following any common distribution from which random samples are taken and you wish to test whether these distributions are equal. Fisher (1993) page 122 and Mardia & Jupp (2000) pages 156-157 present a nonparametric test that is calculated as follows

$$W_g = 2\sum_{i=1}^{g} \frac{\left(C_{Ri}^2 + S_{Ri}^2\right)}{n_i}$$

where $C_{Ri} = \sum_{j=1}^{n_i} \cos(\gamma_{ij})$, $S_{Ri} = \sum_{j=1}^{n_i} \sin(\gamma_{ij})$, $n = \sum_{i=1}^{g} n_i$, and $\gamma_{ij}$ are the circular ranks of the corresponding angles. The circular ranks are calculated using

$$\gamma_{ij} = \frac{2\pi r_{ij}}{n}$$

where the $r_{ij}$ are the ranks of the corresponding $a_{ij}$.

If all $n_i$ are greater than 10, the distribution of $W_g$ is approximately distributed as a chi-square with 2*g*-2 degrees of freedom.

Since ranks are used in this test, ties become an issue. We have adopted the strategy of applying average ranks. Note that little has been done to test the adoption of this strategy within the realm of circular statistics.

### Watson-Williams High Concentration F Test

Suppose you have *g* Von Mises populations from which random samples are taken and you wish to test whether their mean directions are equal. That is, you wish to test the null hypothesis

$$H_0: \theta_1 = \theta_2 = \ldots = \theta_g$$

Mardia & Jupp (2000) pages 134-135 present the Watson-William High-Concentration F Test that is calculated as follows

$$F_{ww} = \left(1 + \frac{3}{8\hat{\kappa}}\right) \frac{\left(\sum_{j=1}^{g} R_j - R\right) / (g-1)}{\left(n - \sum_{j=1}^{g} R_j\right) / (n-g)}$$

where $\hat{\kappa}$ is the maximum likelihood estimate of the concentration based on $R$ and

$R_j = \sqrt{C_j^2 + S_j^2}$, $C_j = \sum_{i=1}^{n_j} \cos(a_i)$, $S_j = \sum_{i=1}^{n_j} \sin(a_i)$, $R = \sqrt{C^2 + S^2}$, $C = \sum_{j=1}^{g} C_j$, $S = \sum_{j=1}^{g} S_j$, and $n = \sum_{j=1}^{g} n_j$.

The distribution of $F_{ww}$ is approximately distributed as an $F$ with $g-1$ and $n-1$ degrees of freedom when the assumptions that $\kappa_1 = \kappa_2 = ... = \kappa_g$ and that the distributions are Von Mises are made. The approximation also requires that $\hat{\kappa} \geq 1$.

## Multi-Group Concentration Homogeneity Test

Suppose you have $g$ groups from which random samples are taken and you wish to test whether the concentrations are equal. That is, you wish to test the null hypothesis

$$H_0: \kappa_1 = \kappa_2 = ... = \kappa_g$$

Mardia & Jupp (2000) page 139 presents such a test. It is divided into three cases.

### Case I. $\overline{R} < 0.45$

$U1$ is approximately distributed as a chi-square with $g-1$ degrees of freedom

$$U1 = \sum_{j=1}^{g} w_j f_j^2 - \frac{\left\{\sum_{j=1}^{g} w_j f_j\right\}^2}{\sum_{j=1}^{g} w_j}$$

where $w_j = \frac{4(n_j - 4)}{3}$ and $f_j = \sin^{-1}\left(2\overline{R}_j\sqrt{3/8}\right)$

### Case II. $0.45 \leq \overline{R} \leq 0.70$

$U2$ is approximately distributed as a chi-square with $g-1$ degrees of freedom

$$U2 = \sum_{j=1}^{g} w_j h_j^2 - \frac{\left\{\sum_{j=1}^{g} w_j h_j\right\}^2}{\sum_{j=1}^{g} w_j}$$

where $w_j = \dfrac{n_j - 3}{0.797449}$ and $h_j = \sinh^{-1}\left(\dfrac{\overline{R}_j - 1.089}{.258}\right)$

**Case III.** $\overline{R} > 0.70$

*U*3 is approximately distributed as a chi-square with *g-1* degrees of freedom

$$U3 = \frac{1}{1+d}\left\{\nu \log\left(\frac{n - \sum_{j=1}^{g} R_j}{\nu}\right) - \sum_{j=1}^{g} \nu_j \log\left(\frac{n_j - R_j}{\nu_j}\right)\right\}$$

where $\nu_j = n_j - 1$, $\nu = n - g$, and $d = \dfrac{1}{3(g-1)}\left\{\sum_{j=1}^{g}\frac{1}{\nu_j} - \frac{1}{\nu}\right\}$.

## Circular Correlation Measure

This section discusses a measure of the correlation between two circular variables presented by Jammalamadaka and SenGupta (2001). Suppose a sample of *n* pairs of angles $(a_{11}, a_{21}), (a_{12}, a_{22}), ..., (a_{1n}, a_{2n})$ is available. The circular correlation coefficient is calculated as

$$r_c = \frac{\sum_{k=1}^{n} \sin(a_{1k} - T_{1,1})\sin(a_{2k} - T_{2,1})}{\sqrt{\sum_{k=1}^{n} \sin^2(a_{1k} - T_{1,1}) \sum_{k=1}^{n} \sin^2(a_{2k} - T_{2,1})}}$$

where $T_{1,1}$ is the mean direction of the first circular variable and $T_{2,1}$ is the mean direction of second.

The significance of this correlation coefficient can be test using the fact the $z_r$ is approximately distributed as a standard normal, where

$$z_r = r_c \sqrt{\frac{n\lambda_{20}\lambda_{02}}{\lambda_{22}}}$$

and

$$\lambda_{ij} = \frac{1}{n}\sum_{k=1}^{n} \sin^i(a_{1k} - T_{1,1})\sin^j(a_{2k} - T_{2,1})$$

## Data Structure

The data consist of one or more variables. Each variable contains a set of angular values. The rows may be separated into groups using the unique values of an optional grouping variable. An example of a dataset containing circular data is CIRCULAR1.S0. Missing values are entered as blanks (empty cells).

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

These options specify the variables that will be used in the analysis.

### Data Variables

#### Data Variables

Specify one or more variables that contain the angular values. The values in these variables must be of the type specified in 'Data Type'.

If more than one variable is specified, the format of the reports depend on whether a 'Grouping Variable' is used. If a 'Grouping Variable' is specified, a separate set of reports is generated for each data variable. If no 'Grouping Variable' is specified, each of these variables are treated as a different group in a single set of reports.

#### Data Type

Specify the type of circular data that is contained in the Data Variables. Note that all variables must be of the same data type. The possible data types are

- **Angle (0 to 360)**

  Data are in the range 0 to 360 degrees. Negative values are converted to positive values by subtracting them from 360 (e.g. -20 becomes 340). Data outside 0 to 360 are converted to this range by subtracting (or adding) 360 until the value is in this range.

- **RADIAN (0 to 2 pi)**

  Data are in the range 0 to 2pi radian. Negative values are converted to positive values by subtracting them from 2pi.

- **AXIAL (0 to 180)**

  Data are bidirectional. Axial data are converted to angular data by multiplying by two. Axial data may be in the full 0-360 range.

- **Compass**

  Text data representing the 16 points of the compass are entered. Values are converted into degrees using the recodes: N = 0, E = 90, S = 180, W = 270. Two and three letters may be used. For example, 'NNW' is north by north-west.

- **Time (0-24)**

  Time of day values between 0 and 24 may be entered.

- **Weekday**

  Integers representing the days of the week are entered. The relationship is 1 = Monday, 2 = Tuesday, ..., 7 = Sunday. The integers are converted to degrees using 1 = 180/7, 2 = 180/7+360/7, and so on.

- **Month of Year**

  Integers representing the months of the year are entered. The relationship is 1 = January, 2 = February, ..., 12 = December. The integers are converted to degrees using 1 = 180/12, 2 = 180/12+360/12, and so on.

## Grouping Variable

### Grouping Variable

This optional variable separates the values of the Data Variables into groups. A separate analysis is then generated for each group.

Note that when a grouping variable is specified, the correlations are not generated.

### Grouping Correction Factor

When the same data values occur repeatedly, a correction factor is suggested for the calculation of R bar. This correction factor depends on the number of unique values, which is entered here. If '0' is entered, no correction factor is used.

## Options – Hypothesized Values

### Hypothesized Theta

This optional parameter specifies the hypothesized value of theta (mean direction) under the null hypothesis. A set of hypotheses tests are conducted to determine if the data support this hypothesized value.

Note that this is a single-group test. If there are several groups, a separate test is provided for each group.

### Hypothesized Kappa

This optional parameter specifies the hypothesized value of kappa (concentration) under the null hypothesis. A hypothesis test is conducted to determine if the data support this hypothesized value.

Note that this is a single-group test. If there are several groups, a separate test is provided for each group.

## Options – Confidence Coefficient

### Confidence Coefficient

Specify the value of confidence coefficient for the confidence intervals.

# Reports Tab

The options on this panel control which reports and plots are displayed.

## Select Reports

### Summary Reports ... Correlations

Select these options to display the indicated reports.

## Select Plots

### Rose Plot (Combined) and Rose Plots (Individual)

Select these options to display the indicated plots.

## Report Options

### Show Notes

This option controls whether the available notes and comments that are displayed at the bottom of each report. This option lets you omit these notes to reduce the length of the output.

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Value Labels

This option applies to the *Group Variable(s)*. It lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

## Report Options – Decimal Places

### Mean and Probability Decimals

Specify the number of digits after the decimal point to display on the output of values of this type. Note that this option in no way influences the accuracy with which the calculations are done.

Enter 'All' to display all digits available. The number of digits displayed by this option is controlled by whether the PRECISION option is SINGLE or DOUBLE.

# Plot Options 1 Tab

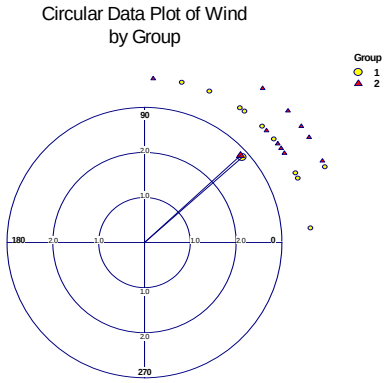
The options on this panel control the appearance of the plots.

## Plot Contents

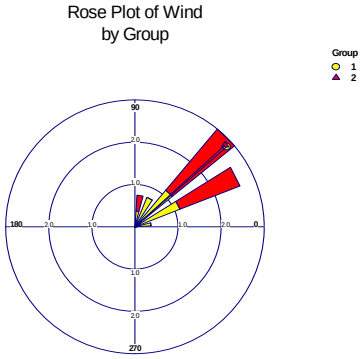
### Objects on Plot

This setting controls which objects are displayed on the plots. The possible settings are

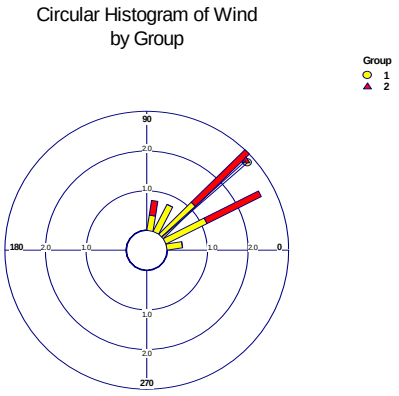
- **Raw Data**



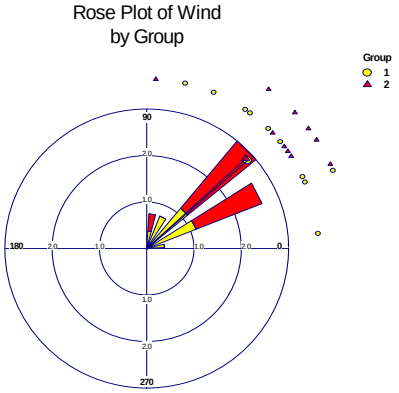

- **Rose Plot**



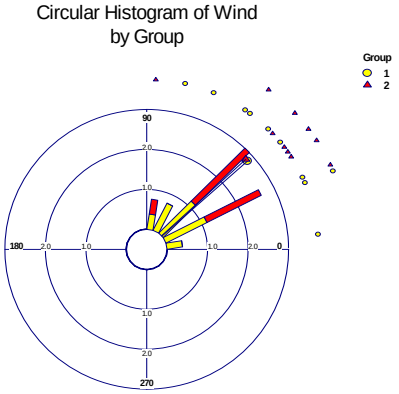
- **Circular Histogram**




- **Raw Data & Rose Plot**




- **Raw Data & Histogram**

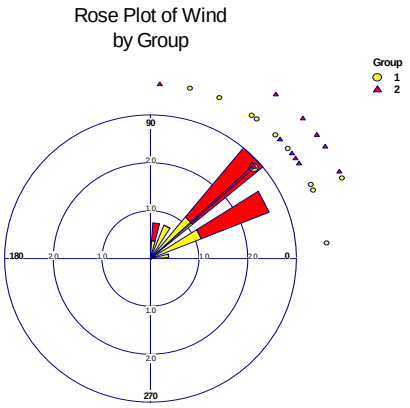


## Display Type

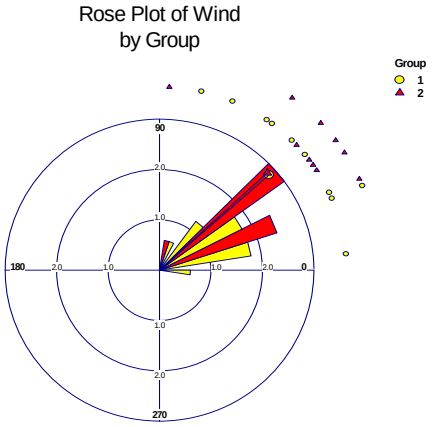
### Group Display Type

Specify whether the group are 'Stacked' or 'Side-by-Side'.

- **Stacked**




- **Side-by-Side**




## Plot Setup

### Data Direction on Plot

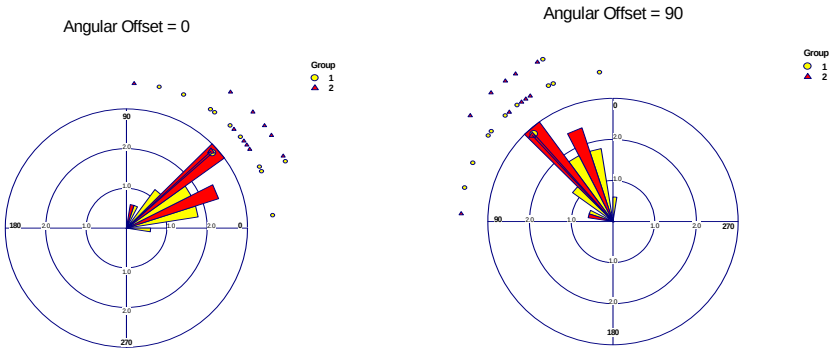
This option indicates whether the orientation of the plot is in a 'Clockwise' or 'Counter-Clockwise' direction.

### Angular Offset on Plot

This option lets you indicate the position of 0 degrees by entering an offset angle. On the default circle, 0 degrees is on the right (east), 90 degrees is at the top (north), 180 degrees is on the left

(west), and 270 degrees is at the bottom (south). This option lets you add an 'offset' to each angle which moves the position of 0 degrees around the circle.

The offset must be between 0 and 360 degrees.




### Histogram and Rose Plot Bins

Specify the number of bins (bars) to be displayed on the circular histogram or rose plot. A reasonable value is 20. This will cause each bin to have a width of 360/20 = 18 degrees.

### Data Bins

Specify the number of positions around the circle at which data values will be plotted. The recommended value is 180.

### Percent Inside Circle

Imagine that that plotting surface is a circle. This parameter sets the percent of the overall radius that is devoted to the rose plot (or histogram). That is, it is the percent of the plot that is inside the circle. 100 minus this amount is the percentage devoted to the plotting of the raw data outside the circle.

### Percent Histogram Base

Imagine that the plotting surface is a circle. This parameter specifies the percent of the radius of this circle that is devoted to the base of the histogram.

A good value is '10'.

### Rose Petal Width

This option is for the rose plot only, when the Group Display Type is set to 'Stacked'. It is the percent of the bin width that is used for the petal. The remaining space is empty (blank).

A good value is '50'.

### Radius of Mean Symbol

Specify the radius of the symbols used to represent the mean directions on the plot. The typical value is 100.

Enter '0' if you do not want the mean displayed on the plot.

## Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, {G} is replaced by the appropriate grouping variable's name.

## Titles – Combined Plot and Individual Plot Titles

### Title Line 1 and Title Line 2

This is the text of the title(s). The characters *{Y}*, *{S}*, and *{G}* are replaced by appropriate variable names, an internal phrase, and the grouping variable's name, respectively. Press the button on the right of the field to specify the font of the text.

# Plot Options 2 Tab

The options on this panel control the appearance of the plots.

## Circular Axes – Outside Circle

These options control the outside axis of the plots.

### Show

Check this option to display the outside (main) circular axis on the plots.

### Axis Line

This option controls the format of the outside circular axis line. Click on the arrow button to the right to edit the settings.

## Circular Axes – Interior Circle(s)

These options control the interior axes of the plots.

### Number

This is the number of circular axes (shown as circles on the plots). The recommended value is 2.

### Axis Line

This option controls the format of the interior circular axis lines. Click on the arrow button to the right to edit the settings.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed along each axis.

## Radial Axes

These options control the radial axes (spokes) of the plots.

### Number

This is the number of radial axes. The recommended value is 4

### Axis Line

This option controls the format of the radial axis lines. Click on the arrow button to the right to edit the settings.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed along each axis.

### Reference Number Offset

The reference numbers of the radial axes may be offset slightly for a better plot. This parameter controls the amount of this offset as a percentage of the overall circular radius. Values near 100 print near the edge of the circle. Values near zero print towards the center of the circle. The recommended value is 92.

## Plot Colors

These options control the colors used in the plots.

### Background and Interior Color

These options specify the plot background and interior colors. Click the button at the right to change the colors.

## Plotting Symbols

These options control the symbols used in the plots.

### Symbol (1-15)

The symbols used to represent the groups. Symbol 1 represents the first group, Variable 2 represents the second group, and so on.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Analysis of Circular Data

This section presents an example of how to run this procedure. The data are wind directions of two groups. The data are found in the CIRCULAR1.S0 database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Circular Data Analysis window.

**1 Open the CIRCULAR1 dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **Circular1.s0**.
- Click **Open**.

**2 Open the Circular Data window.**

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Circular Data Analysis**. The Circular Data Analysis procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Circular Data window, select the **Variables tab**. (This is the default.)
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Wind** from the list of variables and then click **Ok**. "Wind" will appear in the Data Variables box.
- Double-click in the **Grouping Variable** text box. This will bring up the variable selection window.
- Select **Group** from the list of variables and then click **Ok**. "Group" will appear in the Paired Variables box.
- Set **Hypothesized Theta** to **40**.
- Set **Hypothesized Kappa** to **2**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

The following reports and charts will be displayed in the Output window.

## Summary Statistics Section

| Group | Sample Size (N) | Mean Direction (Theta) | Mean Resultant Length (R bar) | Circular Variance (V) | Circular Standard Deviation (v) | Circular Dispersion (Delta) | Von Mises Concentration (Kappa) |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 41.5869 | 0.9324 | 0.0676 | 21.4299 | 0.1449 | 5.5452 |
| 2 | 10 | 42.6725 | 0.9599 | 0.0401 | 16.3991 | 0.0768 | 9.1850 |

### Group

This is the group (or variable) presented on this line.

### Sample Size

This is the number of nonmissing values in this group.

### Mean Direction

This is estimated mean direction, $T_1$.

### Mean Resultant Length

This is the estimated mean resultant length, $\overline{R}_1$. It is a measure of data concentration. An $\overline{R}_1$ close to zero implies low data concentration. An $\overline{R}_1$ close to one implies high data concentration.

### Circular Variance

The circular variance, *V*, is a measure of variation in the data. Note that $V = 1 - \overline{R}_1$.

### Circular Standard Deviation

The circular standard deviation is $v = \sqrt{-2\ln\left(\overline{R}_1\right)}$. Note that it is not the square root of the circular variance.

### Circular Dispersion

The circular dispersion, $\delta = \frac{1 - T_2}{2\overline{R}_1^2}$, is another measure of variation.

### Von Mises Concentration

This is the estimated concentration parameter of the von Mises distribution, $\kappa$.

## Mean Direction Section

| Group | Sample Size (N) | Mean Direction (Theta) | Lower 95.0% Confidence Limit of Theta | Upper 95.0% Confidence Limit of Theta | Standard Error of Mean Direction |
|---|---|---|---|---|---|
| 1 | 10 | 41.5869 | 27.9417 | 55.2321 | 6.8964 |
| 2 | 10 | 42.6725 | 32.7516 | 52.5934 | 5.0365 |

This report provides the large sample confidence interval for the mean direction as described by Upton & Fingleton (1989) page 220. Note that this interval does not require the assumption that the data come from the von Mises distribution.

## Variation Statistics Section

| Group | Sample Size (N) | Circular Variance (V) | Circular Standard Deviation (v) | Circular Dispersion (Delta) | Skewness (s) | Kurtosis (k) |
|---|---|---|---|---|---|---|
| 1 | 10 | 0.0676 | 21.4299 | 0.1449 | -0.0795 | -1.7244 |
| 2 | 10 | 0.0401 | 16.3991 | 0.0768 | -4.8582 | 5.4248 |

This report provides measures of data variation and dispersion which were defined in the Statistical Summary Report. It also provides measures of the skewness and kurtosis of the data.

### Skewness

This is a measure of the skewness (lack of symmetry about the mean) in the data. Symmetric, unimodal datasets have a skewness value near zero.

### Kurtosis

This is a measure of the kurtosis (peakedness) in the data. Von Mises datasets have a kurtosis near zero.

## Von Mises Distribution Estimation Section

| Group | Sample Size (N) | Mean Direction (Theta) | Lower 95.0% Confidence Limit of Theta | Upper 95.0% Confidence Limit of Theta | Von Mises Conc. (Kappa) | Lower 95.0% Confidence Limit of Kappa | Upper 95.0% Confidence Limit of Kappa |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 41.5869 | 29.1135 | 54.0603 | 5.5452 | 2.3222 | 8.4051 |
| 2 | 10 | 42.6725 | 33.2995 | 52.0455 | 9.1850 | 3.0227 | 13.2191 |

This report provides estimates and confidence intervals of the parameters (mean direction and concentration) of the von Mises distribution that best fits the data. Note that the von Mises distribution is a symmetric, unimodal distribution. You should check the rose plot or circular histogram to determine if the data are symmetric.

The formulas used in the estimation and confidence intervals were given earlier in this chapter. They come from Mardia & Jupp (2000).

## Trigonometric Moments Section

| Group | N | Mean Cos(a) | Mean Sin(a) | Mean Cos(2a) | Mean Sin(2a) | R bar | 2R bar | Theta | 2Theta |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 10 | 0.6974 | 0.6189 | 0.0903 | 0.7426 | 0.9324 | 0.7481 | 41.5869 | 83.0670 |
| 2 | 10 | 0.7057 | 0.6506 | 0.1085 | 0.8516 | 0.9599 | 0.8585 | 42.6725 | 82.7373 |

This report provides summary statistics that are used in other calculations.

### Mean Cos(a)

This is $\overline{C_1} = \frac{1}{n}\sum_{i=1}^{n}\cos(a_i)$.

### Mean Sin(a)

This is $\overline{S}_1 = \frac{1}{n}\sum_{i=1}^{n}\sin(a_i)$.

### Mean Cos(2a)

This is $\overline{C}_2 = \frac{1}{n}\sum_{i=1}^{n}\cos(2a_i)$.

### Mean Sin(2a)

This is $\overline{S}_2 = \frac{1}{n}\sum_{i=1}^{n}\sin(2a_i)$.

### R bar

This is $\overline{R}_1 = \frac{1}{n}\sqrt{n\left(\overline{C}_1^2 + \overline{S}_1^2\right)}$.

### 2R bar

This is $\overline{R}_2 = \frac{1}{n}\sqrt{n\left(\overline{C}_2^2 + \overline{S}_2^2\right)}$.

### Theta, 2 Theta

This is calculated using the following formula with *p* set to 1 and then 2, respectively.

$$T_p = \begin{cases} \tan^{-1}\left(\frac{\overline{S}_p}{\overline{C}_p}\right) & \overline{C}_p > 0, \overline{S}_p > 0 \\ \tan^{-1}\left(\frac{\overline{S}_p}{\overline{C}_p}\right) + \pi & \overline{C}_p < 0 \\ \tan^{-1}\left(\frac{\overline{S}_p}{\overline{C}_p}\right) + 2\pi & \overline{S}_p < 0, \overline{C}_p > 0 \end{cases}$$

## Multiple-Group Hypothesis Tests Section

| Null Hypothesis (H0) | Test Name | Test Statistic | Prob Level | Reject H0 at 0.05 Level |
|---|---|---|---|---|
| Equal Distributions | Uniform Scores Test | 6.7392 | 0.0344 | Yes |
| Equal Directions | Watson-Williams F Test | 0.0147 | 0.9047 | No |
| Equal Concentrations | Concentration Homogeneity Test | 0.5717 | 0.4496 | No |

**Notes:**
These statistics test various hypotheses about the parameters of von Mises distributions.
They require that each group follow the von Mises distribution.
The Uniform Scores test requires samples of at least 10.
The Watson-Williams F-test assumes that all kappa's are equal and that their average is > 1.

This report provides tests for three hypotheses about the features of several von Mises datasets. That is, it provides a test of whether the distributions are identical, whether the mean directions

are identical, and whether the concentrations are identical. These tests are documented in the Technical Details section of this chapter.

## Two-Group Hypothesis Tests Section

| | | Equal Distributions | | Equal Directions | | Equal Concentrations | |
|---|---|---|---|---|---|---|---|
| First Group | Second Group | Test Statistic | Prob Level | Test Statistic | Prob Level | Test Statistic | Prob Level |
| 1 | 2 | 6.7392 | 0.0344 | 0.0147 | 0.9047 | 0.5717 | 0.4496 |

**Notes:**
These statistics test various hypotheses about the parameters of von Mises distributions.
They require that each group follow the von Mises distribution.
Equal distributions tested by the Mardia-Watson-Wheeler uniform scores test. Requires all Ni > 10.
Equal directions tested by the Watson-Williams F test. Assumes Von Mises data with equal kappa's, all > 1.
Equal concentrations tested by concentration homogeneity test. Assumes Von Mises data.

This report provides the same three tests as the Multiple-Group Hypothesis Tests Section, taken two groups at a time. It allows you to pinpoint where differences occur.

## Tests for a Specified Mean Direction Assuming Von Mises Data – Test Statistic & Prob Levels

**Tests for a Specified Mean Direction Assuming Von Mises Data - Test Statistics - Wind**

| Group | Sample Size (N) | Actual Mean Direction (Theta) | H0 Mean Direction (Theta0) | Score Test Z Value | Likelihood Ratio CS Value | Watson & Williams F Value | Stephens Test Z Value |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 41.5869 | 40.0000 | 0.0409 | 0.0470 | 0.0476 | 0.0397 |
| 2 | 10 | 42.6725 | 40.0000 | 0.1949 | 0.2266 | 0.2341 | 0.1917 |

**Notes:**
These procedures test whether the mean direction is equal to a specified value, when kappa (concentration) is unknown.
They assume that the data follow the von Mises distribution.
The Score Test requires a large sample size.
The Likelihood Ratio Test requires a sample size of at least 5.
The Watson & Williams Test requires a large value of kappa.
The Stephens Test requires kappa to be greater than 2.

**Tests for a Specified Mean Direction Assuming Von Mises Data - Probability Levels - Wind**

| Group | Sample Size (N) | Actual Mean Direction (Theta) | H0 Mean Direction (Theta0) | Score Test Prob Level | Likelihood Ratio Prob Level | Watson & Williams Prob Level | Stephens Test Prob Level |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 41.5869 | 40.0000 | 0.8398 | 0.8284 | 0.8321 | 0.8422 |
| 2 | 10 | 42.6725 | 40.0000 | 0.6589 | 0.6340 | 0.6400 | 0.6615 |

**Notes:**
This report gives the probability levels of the test statistics displayed in the previous report.
Although the probability levels of four tests are given, you should use only one of these.

This section reports the results of four tests of the hypothesis that the mean direction of a particular group is equal to a specific value. These are two-sided tests. They were documented earlier in this chapter.

The first table gives the values of the test statistics. The second table gives the probability levels. The null hypothesis is rejected when the probability level is less than 0.05 (or some other appropriate cutoff).

## Tests for a Specified Concentration Assuming Von Mises Data

| Group | Sample Size (N) | Actual Concentration (Kappa) | H0 Concentration (Kappa0) | Chi-Square Value | Prob Level of (H1:Kappa < Kappa0) | Prob Level of (H1:Kappa > Kappa0) |
|---|---|---|---|---|---|---|
| 1 | 10 | 5.5452 | 2.0000 | 2.2756 | 0.0137 | 0.9863 |
| 2 | 10 | 9.1850 | 2.0000 | 1.3518 | 0.0019 | 0.9981 |

**Notes:**
These statistics test whether the kappa (concentration) parameter is equal to the specified value.
The tests require that the estimated kappa is > 2.

This section reports the results of two, one-sided tests of the hypothesis that the concentration parameter of each group is equal to a specific value. They were documented earlier in this chapter.

The first probability level is for testing the null hypothesis that kappa is greater than or equal to kappa0. The second probability level is for test the null hypothesis that kappa is less than or equal to kappa0.

## Uniform Distribution Goodness-of-Fit Tests

| Group | Sample Size (N) | Rayleigh's Test Statistic (S*) | Rayleigh's Test Prob Level | Kuiper's Test Statistic (V) | Kuiper's Test Prob Level | Watson's Test Statistic (U2) | Watson's Test Prob Level |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 20.2993 | 0.0000 | 2.7145 | 0.0000 | 0.5657 | 0.0001 |
| 2 | 10 | 21.7499 | 0.0000 | 2.8088 | 0.0000 | 0.6788 | 0.0000 |

**Notes:**
The tests in this report assess the goodness-of-fit of the uniform distribution.
The Rayleigh test requires samples of at least 20.
The Kuiper and Watson tests require samples of at least 8.

This section reports the results of three goodness-of-fit tests for the uniform distribution. They were documented earlier in this chapter.

These tests may be viewed as testing whether the data are distributed uniformly around the circle.

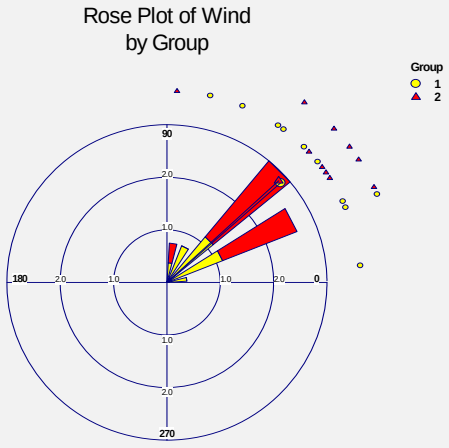
## Von Mises Distribution Goodness-of-Fit Tests

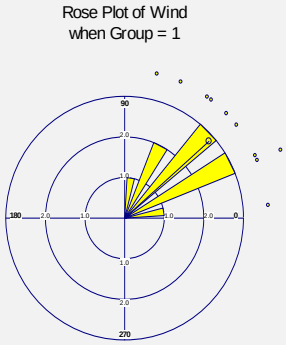
| Group | Sample Size (N) | Watson's Test Statistic (U2) | Watson's Test Prob Level | Cox's Test Statistic (S) | Cox's Test Prob Level |
|---|---|---|---|---|---|
| 1 | 10 | 0.0340 | 0.5000 | 0.4030 | 0.8175 |
| 2 | 10 | 0.1282 | 0.0322 | 2.9309 | 0.2310 |

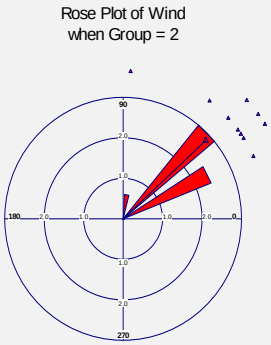
**Notes:**
The tests in this report assess the goodness-of-fit of the von Mises distribution.
Both tests require samples of at least 20.

This section reports the results of two goodness-of-fit tests for the von Mises distribution. They were documented earlier in this chapter. Several hypothesis tests assume that the data follow a von Mises distribution. These tests allow you to check the accuracy of this assumption.

## Rose Plots

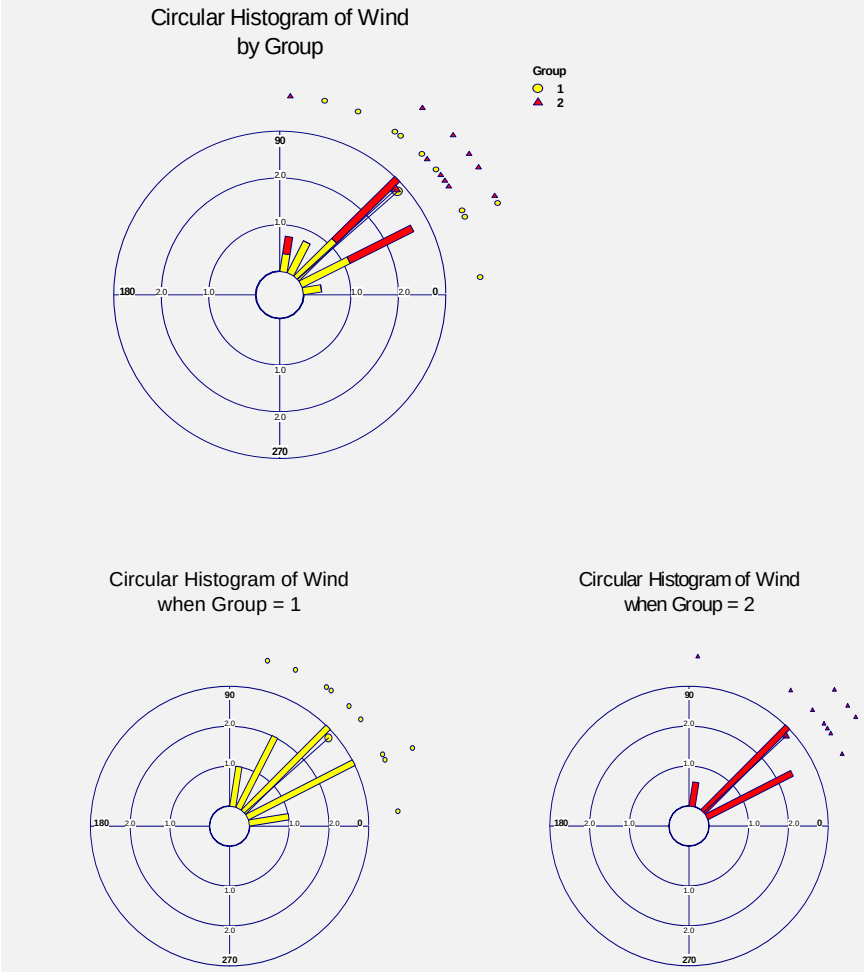



These plots show the distribution of the data around the circle. Although the rose plot is popular, it distorts the counts so that the emphasis is on the larger bins. For this reason, we prefer the circular histograms.

# Circular Histograms




The circular histograms are generated by setting the Objects on Plot to 'Raw Data & Histogram' under the Plot Options tab. Notice that no exact emphasis is placed on the bins with larger counts.

Chapter 235

# Cross-Over Analysis Using T-Tests

## Introduction

This procedure analyzes data from a two-treatment, two-period (2x2) cross-over design. The response is assumed to be a continuous random variable that follows the normal distribution.

In the two-period cross-over design, subjects are randomly assigned to one of two groups. One group receives treatment *R* followed by treatment *T*. The other group receives treatment *T* followed by treatment *R*. Thus, the response is measured at least twice on each subject.

Cross-over designs are used when the treatments alleviate a condition, rather than effect a cure. After the response to one treatment is measured, the treatment is removed and the subject is allowed to return to a baseline response level. Next, the response to a second treatment is measured. Hence, each subject is measured twice, once with each treatment.

Examples of the situations that might use a cross-over design are the comparison of anti-inflammatory drugs in arthritis and the comparison of hypotensive agents in essential hypertension. In both of these cases, symptoms are expected to return to their usual baseline level shortly after the treatment is stopped.

## Equivalence

Cross-over designs are popular in the assessment of equivalence. In this case, the effectiveness of a new treatment formulation (drug) is to be compared against the effectiveness of the currently used (reference) formulation. When showing equivalence, it is not necessary to show that the new treatment is better than the current treatment. Rather, the new treatment need only be shown to be as good as the reference so that it can be used in its place.

## Advantages of Cross-Over Designs

A comparison of treatments on the same subject is expected to be more precise. The increased precision often translates into a smaller sample size. Also, patient enrollment into the study may be easier because each patient will receive both treatments.

## Disadvantages of Cross-Over Designs

The statistical analysis of a cross-over experiment is more complex than a parallel-group experiment and requires additional assumptions. It may be difficult to separate the treatment effect from the time effect and the carry-over effect of the previous treatment.

The design cannot be used when the treatment (or the measurement of the response) alters the subject permanently. Hence, it cannot be used to compare treatments that are intended to effect a cure.

Because subjects must be measured at least twice, it may be more difficult to keep patients enrolled in the study. It is arguably simpler to measure a subject once than to obtain their measurement twice. This is particularly true when the measurement process is painful, uncomfortable, embarrassing, or time consuming.

# Technical Details

## Cross-Over Analysis

In the discussion that follows, we summarize the presentation of Chow and Liu (1999). We suggest that you review their book for a more detailed presentation.

The general linear model for the standard 2x2 cross-over design is

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{(j,k)} + C_{(j-1,k)} + e_{ijk}$$

where *i* represents a subject (1 to $n_k$), *j* represents the period (1 or 2), and *k* represents the sequence (1 or 2). The $S_{ik}$ represent the random effects of the subjects. The $P_j$ represent the effects of the two periods. The $F_{(j,k)}$ represent the effects of the two formulations (treatments). In the case of the 2x2 cross-over design

$$F_{(j,k)} = \begin{cases} F_R & if\ k = j \\ F_T & if\ k \neq j \end{cases}$$

where the subscripts *R* and *T* represent the *reference* and *treatment* formulations, respectively.

The $C_{(j-1,k)}$ represent the carry-over effects. In the case of the 2x2 cross-over design

$$C_{(j-,k)} = \begin{cases} C_R & if\ j = 2, k = 1 \\ C_T & if\ j = 2, k = 2 \\ 0 & otherwise \end{cases}$$

where the subscripts *R* and *T* represent the *reference* and *treatment* formulations, respectively.

Assuming that the average effect of the subjects is zero, the four means from the 2x2 cross-over design can be summarized using the following table.

| *Sequence* | *Period* 1 | *Period* 2 |
|---|---|---|
| 1 (*RT*) | $\mu_{11} = \mu + P_1 + F_R$ | $\mu_{21} = \mu + P_2 + F_T + C_R$ |
| 2 (*TR*) | $\mu_{12} = \mu + P_1 + F_T$ | $\mu_{22} = \mu + P_2 + F_R + C_T$ |

where $P_1 + P_2 = 0$, $F_T + F_R = 0$, and $C_T + C_R = 0$.

## Carryover Effect

The 2x2 cross-over design should only be used when there is no carryover effect from one period to the next. The presence of a carryover effect can be studied by testing whether $C_T = C_R = 0$ using a *t* test. This test is calculated as follows

$$T_c = \frac{\hat{C}}{\hat{\sigma}_u \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$\hat{C} = \overline{U}_{\cdot 2} - \overline{U}_{\cdot 1}$$

$$\overline{U}_{\cdot k} = \frac{1}{n_k} \sum_{i=1}^{n_k} U_{ik}$$

$$\hat{\sigma}_u^2 = \frac{1}{(n_1 - n_2 - 2)} \sum_{k=1}^{2} \sum_{i=1}^{n_k} \left( U_{ik} - \overline{U}_{\cdot k} \right)^2$$

$$U_{ik} = Y_{i1k} + Y_{i2k}$$

The null hypothesis of no carryover effect is rejected at the $\alpha$ significance level if

$$|T_c| > t_{\alpha/2, n_1 + n_2 - 2}.$$

A $100(1-\alpha)\%$ confidence interval for $C = C_T - C_R$ is given by

$$\hat{C} \pm \left( t_{\alpha/2, n_1 + n_2 - 2} \right) \hat{\sigma}_u \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

## Treatment Effect

The presence of a treatment (drug) effect can be studied by testing whether $F_T = F_R = 0$ using a *t* test. This test is calculated as follows

$$T_d = \frac{\hat{F}}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$\hat{F} = \bar{d}_{\cdot 1} - \bar{d}_{\cdot 2}$$

$$\bar{d}_{\cdot k} = \frac{1}{n_k}\sum_{i=1}^{n_k} d_{ik}$$

$$\hat{\sigma}_d^2 = \frac{1}{(n_1 - n_2 - 2)}\sum_{k=1}^{2}\sum_{i=1}^{n_k}\left(d_{ik} - \bar{d}_{\cdot k}\right)^2$$

$$d_{ik} = \frac{Y_{i2k} - Y_{i1k}}{2}$$

The null hypothesis of no drug effect is rejected at the $\alpha$ significance level if

$$|T_d| > t_{\alpha/2, n_1+n_2-2}.$$

A $100(1-\alpha)\%$ confidence interval for $F = F_T - F_R$ is given by

$$\hat{F} \pm \left(t_{\alpha/2, n_1+n_2-2}\right)\hat{\sigma}_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

## Period Effect

The presence of a period effect can be studied by testing whether $P_1 = P_2 = 0$ using a *t* test. This test is calculated as follows

$$T_P = \frac{\hat{P}}{\hat{\sigma}_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$\hat{P} = \bar{O}_{\cdot 1} - \bar{O}_{\cdot 2}$$

$$\bar{O}_{\cdot 1} = \bar{d}_1$$

$$\bar{O}_{\cdot 2} = -\bar{d}_2$$

$$\hat{\sigma}_d^2 = \frac{1}{(n_1 - n_2 - 2)}\sum_{k=1}^{2}\sum_{i=1}^{n_k}\left(d_{ik} - \bar{d}_{\cdot k}\right)^2$$

$$d_{ik} = \frac{Y_{i2k} - Y_{i1k}}{2}$$

The null hypothesis of no drug effect is rejected at the $\alpha$ significance level if

$$|T_P| > t_{\alpha/2, n_1+n_2-2}.$$

A $100(1-\alpha)\%$ confidence interval for $P = P_2 - P_1$ is given by

$$\hat{P} \pm \left(t_{\alpha/2, n_1+n_2-2}\right)\hat{\sigma}_d \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\,.$$

## Bioequivalence

The *t* test of formulations (treatments) may be thought of as a preliminary assessment of bioequivalence. However, this *t* test investigates whether the two treatments are different. It does not assess whether the two treatments are the same—bioequivalent. That is, failure to reject the hypothesis of equal means does not imply bioequivalence. In order to establish bioequivalence, different statistical tests must be used.

Before discussing these tests, it is important to understand that, unlike most statistical hypothesis tests, when testing bioequivalence, you want to establish that the response to the two treatments is the same. Hence, the null hypothesis is that the mean responses are different and the alternative hypothesis is that the mean responses are equal. This is just the opposite from the usual *t* test. This is why bioequivalence testing requires the special statistical techniques discussed here.

When using a cross-over design to test for bioequivalence, a washout period between the first and second periods must be used that is long enough to eliminate the residual effects of the first treatment from the response to the second treatment. Because of this washout period, there is no carryover effect. Without a carryover effect, the general linear model reduces to

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{(j,k)} + e_{ijk}$$

There are many types of bioequivalence. The 2x2 cross-over design is used to assess *average bioequivalence.* Remember that average bioequivalence is a statement about the population average. It does not make reference to the variability in responses to the two treatments. The 1992 FDA guidance uses the $\pm 20$ rule which allows an average response to a test formulation to vary up to 20% from the average response of the reference formulation. This rule requires that ratio of the two averages $\mu_T / \mu_R$ be between 0.8 and 1.2 (80% to 120%). Another way of stating this is that the $\mu_T$ is within 20% of $\mu_R$. The FDA requires that the significance level be 0.10 or less.

Several methods have been proposed to test for bioequivalence. Although the program provides several methods, you should select only the one that is most appropriate for your work.

### Confidence Interval Approach

The confidence interval approach, first suggested by Westlake (1981), states that bioequivalence may be concluded if a $(1-2\alpha)\times 100\%$ confidence interval for the difference $\mu_T - \mu_R$ or ratio $\mu_T / \mu_R$ is within acceptance limits ($\alpha$ is usually set to 0.05). If the $\pm 20$ rule is used, this means that the confidence interval for the difference must be between -0.2 and 0.2. Likewise, the confidence interval for the ratio must be between 0.8 and 1.2 (or 80% and 120%). Several methods have been suggested for computing the above confidence interval. The program provides the results for five of these. Perhaps the best of the five is the one based on Fieller's Theorem since it makes the fewest, and most general, assumptions about the distribution of the responses.

### Classic (Shortest) Confidence Interval of the Difference

$$L_1 = \left(\overline{Y}_T - \overline{Y}_R\right) - \left(t_{\alpha,n_1+n_2-2}\right)\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$U_1 = \left(\overline{Y}_T - \overline{Y}_R\right) + \left(t_{\alpha,n_1+n_2-2}\right)\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

### Classic (Shortest) Confidence Interval of the Ratio

A confidence interval for the ratio may be calculated from the confidence interval on the difference using the formula

$$L_2 = \left(L_1 / \overline{Y}_R + 1\right) \times 100\%$$

$$U_2 = \left(U_1 / \overline{Y}_R + 1\right) \times 100\%$$

### Westlake's Symmetric Confidence Interval of the Difference

First, compute values of $k_1$ and $k_2$ so that

$$1 - 2\alpha = \int_{k_2}^{k_1} T_{n_1+n_2-2}\, dt$$

Next, compute $\Delta$ using

$$\Delta = k_1\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} - \left(\overline{Y}_R - \overline{Y}_T\right)$$

$$= -k_2\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} + 2\left(\overline{Y}_R - \overline{Y}_T\right)$$

Finally, conclude bioequivalence if

$$\left|\Delta\right| < 0.2\mu_R$$

### Westlake's Symmetric Confidence Interval of the Ratio

A confidence interval for the ratio may be calculated from the confidence interval on the difference using the formula

$$L_4 = \left(-\left|\Delta\right| / \overline{Y}_R + 1\right) \times 100\%$$

$$U_4 = \left(\left|\Delta\right| / \overline{Y}_R + 1\right) \times 100\%$$

### Confidence Interval of the Ratio Based on Fieller's Theorem

Both the classic and Westlake's confidence interval for the ratio do not take into account the variability of $\overline{Y}_R$ and the correlation between $\overline{Y}_R$ and $\overline{Y}_T - \overline{Y}_R$. Locke (1984) provides formulas using Fieller's theorem that does take into account the variability of $\overline{Y}_R$. This confidence interval is popular not only because it takes into account the variability of $\overline{Y}_R$, but also the intersubject

variability. Also, it only assumes that the data are normal, but not that the group variances are equal as do the other two approaches.

The $(1-2\alpha)\times 100\%$ confidence limits for $\delta = \mu_T / \mu_R$ are the roots of the quadratic equation

$$\left(\bar{Y}_T - \delta\bar{Y}_R\right)^2 - \left(t_{\alpha,n_1+n_2-2}\right)^2 \omega\left(S_{TT} - 2\delta S_{TR} + \delta^2 S_{RR}\right) = 0$$

where

$$\omega = \frac{1}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$S_{RR} = \frac{1}{(n_1 - n_2 - 2)}\left[\sum_{i=1}^{n_1}\left(Y_{i11} - \bar{Y}_{\cdot 11}\right)^2 + \sum_{i=1}^{n_2}\left(Y_{i22} - \bar{Y}_{\cdot 22}\right)^2\right]$$

$$S_{TT} = \frac{1}{(n_1 - n_2 - 2)}\left[\sum_{i=1}^{n_1}\left(Y_{i21} - \bar{Y}_{\cdot 21}\right)^2 + \sum_{i=1}^{n_2}\left(Y_{i12} - \bar{Y}_{\cdot 12}\right)^2\right]$$

$$S_{TR} = \frac{1}{(n_1 - n_2 - 2)}\left[\sum_{i=1}^{n_1}\left(Y_{i11} - \bar{Y}_{\cdot 11}\right)\left(Y_{i21} - \bar{Y}_{\cdot 21}\right) + \sum_{i=1}^{n_2}\left(Y_{i12} - \bar{Y}_{\cdot 12}\right)\left(Y_{i22} - \bar{Y}_{\cdot 22}\right)\right]$$

Additionally, in order for the roots of the quadratic equation to be finite positive real numbers, the above values must obey the conditions

$$\frac{\bar{Y}_R}{\sqrt{\omega S_{RR}}} > t_{\alpha,n_1+n_2-2}$$

and

$$\frac{\bar{Y}_T}{\sqrt{\omega S_{TT}}} > t_{\alpha,n_1+n_2-2}$$

## Interval Hypotheses Testing Approach

Schuirmann (1981) introduced the idea of using an interval hypothesis to test for average bioequivalence using the following null and alternative hypotheses

$$H_0: \mu_T - \mu_R \le \theta_L \quad or \quad \mu_T - \mu_R \ge \theta_U$$

$$H_a: \theta_L < \mu_T - \mu_R < \theta_U$$

where $\theta_L$ and $\theta_U$ are limits selected to insure bioequivalence. Often these limits are set at 20% of the reference mean. These hypotheses can be rearranged into two one-sided hypotheses as follows

$$H_{01}: \mu_T - \mu_R \le \theta_L \quad versus \quad H_{a1}: \mu_T - \mu_R > \theta_L$$

$$H_{02}: \mu_T - \mu_R \ge \theta_U \quad versus \quad H_{a2}: \mu_T - \mu_R < \theta_U$$

The first hypothesis test whether the treatment response is too low and the second tests whether the treatment response is too high. If both null hypotheses are rejected, you conclude that the treatment drug is bioequivalent to the reference drug.

### Schuirmann's Two One-Sided Tests Procedure

Schuirmann's procedure is to conduct two one-sided tests, each at a significance level of $\alpha$. If both tests are rejected, the conclusion of bioequivalence is made at the $\alpha$ significance level. That is, you conclude that $\mu_T$ and $\mu_R$ are average equivalent at the $\alpha$ significance level if

$$T_L = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \theta_L}{\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} > t_{\alpha,n_1+n_2-2}$$

and

$$T_U = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \theta_U}{\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} > -t_{\alpha,n_1+n_2-2}$$

### Wilcoxon-Mann-Whitney Two One-Sided Tests Procedure

When the normality assumption is suspect, you can use the nonparametric version of Schuirmann's procedure, known as the Wilcoxon-Mann-Whitney two one-sided tests procedure. This rather complicated procedure is described on pages 110 - 115 of Chow and Liu (1999) and we will not repeat their presentation here.

### Anderson and Hauck's Test

Unlike Schuirman's test, Anderson and Hauck (1983) proposed a single procedure that evaluates the null hypothesis of inequivalence versus the alternative hypothesis of equivalence. The significance level of the Anderson and Hauck test is given by

$$\alpha = \Pr\left(|t_{AH}| - \hat{\delta}\right) - \Pr\left(-|t_{AH}| - \hat{\delta}\right)$$

where

$$\Pr(x) = \int_{-\infty}^{x} t_{n_1+n_2-2}\, dt$$

$$\hat{\delta} = \frac{\theta_U - \theta_L}{\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$t_{AH} = \frac{\left(\overline{Y}_T - \overline{Y}_R\right) - \left(\theta_U + \theta_L\right)/2}{\hat{\sigma}_d\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

# Data Structure

The data for a cross-over design is entered into three variables. The first variable contains the sequence number, the second variable contains the response in the first period, and the third variable contains the response in the second period. Note that each row of data represents the complete response for a single subject.

Chow and Liu (1999) give the following data on page 73. We will use these data in our examples to verify the accuracy of our calculations. These data are contained in the database called ChowLiu73.S0.

**CHOWLIU73 dataset**

| Sequence | Period 1 | Period 2 |
|---|---|---|
| 1 | 74.675 | 73.675 |
| 1 | 96.400 | 93.250 |
| 1 | 101.950 | 102.125 |
| 1 | 79.050 | 69.450 |
| 1 | 79.050 | 69.025 |
| 1 | 85.950 | 68.700 |
| 1 | 69.725 | 59.425 |
| 1 | 86.275 | 76.125 |
| 1 | 112.675 | 114.875 |
| 1 | 99.525 | 116.250 |
| 1 | 89.425 | 64.175 |
| 1 | 55.175 | 74.575 |
| 2 | 74.825 | 37.350 |
| 2 | 86.875 | 51.925 |
| 2 | 81.675 | 72.175 |
| 2 | 92.700 | 77.500 |
| 2 | 50.450 | 71.875 |
| 2 | 66.125 | 94.025 |
| 2 | 122.450 | 124.975 |
| 2 | 99.075 | 85.225 |
| 2 | 86.350 | 95.925 |
| 2 | 49.925 | 67.100 |
| 2 | 42.700 | 59.425 |

# Validation

Chow and Liu (1999) use the above dataset throughout their book. Except for some obvious typographical errors that exist in their book, our results match their results exactly. We have also tested the algorithm against examples in other texts. In all cases, ***NCSS*** matches the published results.

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

Following is a list of the procedure's options.

## Variables Tab

The options on this panel specify which variables to use.

### Sequence Variable

#### Sequence Group Variable

Specify the variable containing the sequence number. The values in this column should be either 1 (for the first sequence) or 2 (for the second sequence).

In the case of a bioequivalence study, the program assumes that the reference drug is administered first in sequence 1 and second in sequence 2.

### Period Variables

#### Period 1 Variable

Specify the variable containing the responses for the first period of the cross-over trial, one subject per row.

#### Period 2 Variable

Specify the variable containing the responses for the second period of the cross-over trial, one subject per row.

### Treatment Labels

#### Label 1

This is the one-letter label given to the first treatment. This identifies the treatment that occurs first in sequence 1. In an equivalence trial, this is the label of the reference formulation. Common choices are *R* or *A*.

#### Label 2

This is the one-letter label given to the second treatment. This identifies the treatment that occurs second in sequence 1. In an equivalence trial, this is the label of the treatment formulation. Common choices are *T* or *B*.

### Alpha Levels

#### Cross-Over Alpha Level

This is the value of alpha used in the cross-over reports. One minus alpha is the confidence level of the confidence intervals in the cross-over reports. For example, setting alpha to 0.05 results in a 95% confidence interval.

A value of 0.05 is commonly used. For the preliminary tests, using 0.10 is common. You should not be afraid to use other values since 0.05 became popular in pre-computer days when it was the only value available. Typical values range from 0.001 to 0.20.

### Equivalence Alpha Level

This is the value of alpha used in the equivalence reports. One minus alpha is the confidence level of the confidence intervals in the equivalence reports. For example, setting alpha to 0.05 results in a 95% confidence interval.

You should not be afraid to use values other than 0.05 since this value became popular in pre-computer days when it was the only value available. Typical values range from 0.001 to 0.20.

## Equivalence Limits

### Upper Equivalence Limit

Specify the upper limit of the range of equivalence. Differences between the two treatment means greater than this amount are considered to be bioinequivalent. Note that this should be a positive number.

If the % box is checked, this value is assumed to be a percentage of the reference mean. If the % box is not checked, this value is assumed to be the value of the difference.

### Lower Equivalence Limit

Specify the lower limit of the range of equivalence. Differences between the two treatment means less than this amount are considered to be bioinequivalent. Note that this should be a negative number.

If the % box is checked, this value is assumed to be a percentage of the reference mean. If the % box is not checked, this value is assumed to be the value of the difference.

If you want symmetric limits, enter “-UPPER LIMIT” here and the negative of the Upper Equivalence Limit will be used.

# Reports Tab

The options on this panel control the reports and plots.

## Select Reports

### Cross-Over Summary Report … Written Explanations

Each of these options indicates whether to display the indicated reports.

### Written Explanations

Indicate whether to display the written explanations and interpretations that can be displayed following each report and plot.

## Select Plots

### Means Plot … Probability Plots

Each of these options indicates whether to display the indicated plots.

## Report Options

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Value Labels

This option applies to the *Group Variable(s)*. It lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

## Report Options – Decimal Places

### Mean ... Test Decimals

Specify the number of digits after the decimal point to display on the output of values of this type. Note that this option in no way influences the accuracy with which the calculations are done.

# Means Plot to Period Plot Tabs

The options on this panel control the appearance of various plots.

## Vertical and Horizontal Axis

### Label

This is the text of the axis labels. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Treatments and Connect Subjects

This option lets you specify whether you want to connect the points with a line.

## Plot Settings – Legend

### Show Legend

Indicate whether the legend is to be displayed.

### Legend Text

Indicate the title text of the legend. Note that if two factors are being plotted, {G} is replaced by the appropriate grouping variable's name.

## Titles

### Plot Title

This option contains the text of the plot title. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

# Probability Plot Tab

The options on this panel control the appearance of the probability plot.

## Vertical and Horizontal Axis

### Label

This is the text of the axis labels. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tickmarks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

### Plot Settings

#### Plot Style File

Designate a probability plot style file. This file sets all probability plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Probability Plot procedure.

#### Symbol

Click this box to bring up the symbol specification dialog box. This window will let you set the symbol type, size, and color.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the name of the variable. Press the button on the right of the field to specify the font of the text.

## Symbols Tab

### Plotting Symbols

#### Subject (1-15)

The symbols used to represent the subjects on the Profile Plot. Subject 1 represents the first subject, Subject 2 represents the second subject, and so on.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

#### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Cross-Over Analysis and Validation

This section presents an example of how to run an analysis of data from a 2x2 cross-over design. Chow and Liu (1999) page 73 provide an example of data from a 2x2 cross-over design. These data were shown in the Data Structure section earlier in this chapter. On page 77, they provide the following summary of the results of their analysis.

| Effect | MVUE | Variance | 95% CI | T | P-Value |
|---|---|---|---|---|---|
| Carryover | -9.59 | 245.63 | (-42.10, 22.91) | -0.612 | 0.5468 |
| Treatment | -2.29 | 13.97 | (-10.03, 5.46) | -0.613 | 0.5463 |
| Period | -1.73 | 13.97 | (-9.47, 6.01) | -0.464 | 0.6474 |

We will use the data found the CHOWLIU73 database. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Cross-Over Analysis Using T-Tests window.

**1 Open the CHOWLIU73 dataset.**

- From the **File** menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **CHOWLIU73.s0**.
- Click **Open**.

**2 Open the Cross-Over Analysis Using T-Tests window.**

- On the menus, select **Analysis**, then **T-Tests**, then **Cross-Over Analysis Using T-Tests**. The Cross-Over Analysis Using T-Tests procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Cross-Over Analysis Using T-Tests window, select the **Variables tab**.
- Double-click in the **Sequence Group Variable** box. This will bring up the variable selection window.
- Select **Sequence** from the list of variables and then click **Ok**. The phrase "Sequence" will appear in the Period 2 Variable box.
- Double-click in the **Period 1 Variable** box. This will bring up the variable selection window.
- Select **Period1** from the list of variables and then click **Ok**. The phrase "Period1" will appear in the Period 1 Variable box. Remember that you could have entered a "2" here signifying the second variable on the dataset.
- Double-click in the **Period 2 Variable** box. This will bring up the variable selection window.
- Select **Period2** from the list of variables and then click **Ok**. The phrase "Period2" will appear in the Period 2 Variable box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top of the window).

The following reports and charts will be displayed in the Output window.

## Cross-Over Analysis Summary Section

| Parameter | Estimated Effect | Standard Error | T Value (DF=22) | Prob Level | Lower 95.0% Confidence Limit | Upper 95.0% Confidence Limit |
|---|---|---|---|---|---|---|
| Treatment | -2.29 | 3.73 | -0.61 | 0.5463 | -10.03 | 5.45 |
| Period | -1.73 | 3.73 | -0.46 | 0.6474 | -9.47 | 6.01 |
| Carryover | -9.59 | 15.67 | -0.61 | 0.5468 | -42.09 | 22.91 |

**Interpretation of the Above Report**
The two treatment means in a 2x2 cross-over study are not significantly different at the 0.0500 significance level (the actual significance level was 0.5463). The design had 12 subjects in sequence 1 (RT) and 12 subjects in sequence 2 (TR). The average response to treatment R was 82.56 and the average response to treatment T was 80.27.

A preliminary test failed to reject the assumption of equal period effects at the 0.0500 significance level (the actual significance level was 0.6474). A preliminary test failed to reject the assumption of equal carryover effects at the 0.0500 significance level (the actual significance level was 0.5468).

This report summarizes the results of the analysis. The *Treatment* line presents the results of the t-test of whether the treatments are different. The *Period* line presents the results of a preliminary test of the assumption that the period effects are equal. The *Carryover* line presents the results of a preliminary test of the assumption that there is no carryover effect. This is a critical assumption. If the carryover effect is significant, you should not be using a cross over design.

Note that the values in this report match the values from page 77 of Chow and Liu which validates this part of the program.

### Parameter

These are the items being tested. Note that the *Treatment* line is the main focus of the analysis. The *Period* and *Carryover* lines are preliminary tests of assumptions.

### Estimated Effect

These are the estimated values of the corresponding effects. Formulas for the three effects were given in the Technical Details section earlier in this chapter.

### Standard Error

These are the standard errors of each of the effects. They provide an estimate of the precision of the effect estimate. The formulas were given earlier in the Technical Details section of this chapter.

### T Value (DF=xx)

These are the test statistics calculated from the data that are used to test whether the effect is different from zero.

The *DF* is the value of the degrees of freedom. This is two less than the total number of subjects in the study.

### Prob Level

This is the probability level (p-value) of the test. If this value is less than the chosen significance level, then the corresponding effect is said to be significant. For example, if you are testing at a

significance level of 0.05, then probabilities that are less than 0.05 are statistically significant. You should choose a value appropriate for your study.

Some authors recommend that the tests of assumptions (Period and Carryover) should be done at the 0.10 level of significance.

### Upper and Lower Confidence Limits

These values provide a $(1-\alpha)\times 100\%$ confidence interval for the estimated effect.

### Interpretation of the Above Report

This section provides a written interpretation of the above report.

## Cross-Over Analysis Detail Section

| Seq. | Period | Treatment | Count | Least Squares Mean | Standard Deviation | Standard Error |
|---|---|---|---|---|---|---|
| 1 | 1 | R | 12 | 85.82 | 15.69 | 4.53 |
| 2 | 2 | R | 12 | 79.30 | 25.20 | 7.27 |
| 1 | 2 | T | 12 | 81.80 | 19.71 | 5.69 |
| 2 | 1 | T | 12 | 78.74 | 23.21 | 6.70 |
| 1 | Difference | (T-R)/2 | 12 | -2.01 | 6.42 | 1.85 |
| 2 | Difference | (T-R)/2 | 12 | 0.28 | 11.22 | 3.24 |
| 1 | Total | R+T | 12 | 167.63 | 33.23 | 9.59 |
| 2 | Total | R+T | 12 | 158.04 | 42.93 | 12.39 |
| . | . | R | 24 | 82.56 | | 4.28 |
| . | . | T | 24 | 80.27 | | 4.39 |
| 1 | . | . | 24 | 83.81 | | |
| 2 | . | . | 24 | 79.02 | | |
| . | 1 | . | 24 | 82.28 | | 4.04 |
| . | 2 | . | 24 | 80.55 | | 4.62 |

**Interpretation of the Above Report**
This report shows the means and standard deviations of various subgroups of the data. The least squares mean of treatment R is 82.56 and of treatment T is 80.27. Note that least squares means are created by taking the simple average of their component means, not by taking the average of the raw data. For example, if the mean of the 20 subjects in period 1 sequence 1 is 50.0 and the mean of the 10 subjects in period 2 sequence 2 is 40.0, the least squares mean is (50.0 + 40.0)/2 = 45.0. That is, no adjustment is made for the unequal sample sizes. Also note that the standard deviation and standard error of some of the subgroups are not calculated.

This report provides the least squares means of various subgroups of the data.

### Seq.

This is the sequence number of the mean shown on the line. When the dot (period) appears in this line, the results displayed are created by taking the simple average of the appropriate means of the two sequences.

### Period

This is the period number of the mean shown on the line. When the dot (period) appears in this line, the results displayed are created by taking the simple average of the appropriate means of the two periods.

### Treatment

This is the treatment (or formulation) of the mean shown on the line. When the dot (period) appears in this line, the results displayed are created by taking the simple average of the appropriate means of the two treatments.

When the entry is (*T*-*R*)/2, the mean is computed on the quantities created by dividing the difference in each subject's two scores by 2. When the entry is *R*+*T*, the mean is computed on the sums of the subjects two scores.

### Count

The count is the number of subjects in the mean.

### Least Squares Mean

Least squares means are created by taking the simple average of their component means, not by taking a weighted average based on the sample size in each component. For example, if the mean of the 20 subjects in period 1 sequence 1 is 50.0 and the mean of the 10 subjects in period 2 sequence 2 is 40.0, the least squares mean is (50.0 + 40.0)/2 = 45.0. That is, no adjustment is made for the unequal sample sizes. Since least squares means are used in all subsequent calculations, these are the means that are reported.

### Standard Deviation

This is the estimated standard deviation of the subjects in the mean.

### Standard Error

This is the estimated standard error of the least squares mean.

## Equivalence Based on the Confidence Interval of the Difference

**Equivalence Based on the Confidence Interval of the Difference**

| Test Type | Lower Equivalence Limit | Lower 90.0% Confidence Limit | Upper 90.0% Confidence Limit | Upper Equivalence Limit | Equivalent at the 5.0% Sign. Level? |
|---|---|---|---|---|---|
| Shortest C.I. | -16.51 | -8.70 | 4.12 | 16.51 | Yes |
| Westlake C.I. | -16.51 | -7.41 | 7.41 | 16.51 | Yes |

Note: Westlake's k2 = -1.37 and k1 = 2.60.

**Interpretation of the Above Report**

Average bioequivalence of the two treatments has been found at the 0.0500 significance level using the shortest confidence interval of the difference approach since both confidence limits, -8.70 and 4.12, are between the acceptance limits of -16.51 and 16.51. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

Average bioequivalence of the two treatments has been found at the 0.0500 significance level using Westlake's confidence interval of the difference approach since both confidence limits, -7.41 and 7.41, are between the acceptance limits of -16.51 and 16.51. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

This report provides the results of two tests for bioequivalence based on confidence limits of the difference between the means of the two formulations.

### Test Type

This is the type of test reported on this line. The mathematical details of each test were described earlier in the Technical Details section of this chapter.

### Lower and Upper Equivalence Limit

These are the limits on bioequivalence. As long as the difference between the treatment formulation and reference formula is inside these limits, the treatment formulation is bioequivalent. These values were set by you. They are not calculated from the data.

### Lower and Upper Confidence Limits

These are the confidence limits on the difference in response to the two formulations computed from the data. Note that the confidence coefficient is $(1-2\alpha)\times 100\%$. If both of these limits are inside the two equivalence limits, the treatment formulation is bioequivalent to the reference formulation. Otherwise, it is not.

### Equivalent at the 5.0% Sign. Level?

This column indicates whether bioequivalence can be concluded.

## Equivalence Based on the Confidence Interval of the Ratio

| Test Type | Lower Equivalence Limit | Lower 90.0% Confidence Limit | Upper 90.0% Confidence Limit | Upper Equivalence Limit | Equivalent at the 5.0% Sign. Level? |
|---|---|---|---|---|---|
| Shortest C.I. | 80.00 | 89.46 | 104.99 | 120.00 | Yes |
| Westlake C.I. | 80.00 | 91.02 | 108.98 | 120.00 | Yes |
| Fieller's C.I. | 80.00 | 90.06 | 104.92 | 120.00 | Yes |

**Interpretation of the Above Report**

Average bioequivalence of the two treatments has been found at the 0.0500 significance level using the shortest confidence interval of the ratio approach since both confidence limits, 89.46 and 104.99, are between the acceptance limits of 80.00 and 120.00. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

Average bioequivalence of the two treatments has been found at the 0.0500 significance level using Westlake's confidence interval of the ratio approach since both confidence limits, 91.02 and 108.98, are between the acceptance limits of 80.00 and 120.00. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

Average bioequivalence of the two treatments has been found at the 0.0500 significance level using Fieller's confidence interval of the ratio approach since both confidence limits, 90.06 and 104.92, are between the acceptance limits of 80.00 and 120.00. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

This report provides the results of three tests for bioequivalence based on confidence limits of the ratio of the mean responses to the two formulations.

### Test Type

This is the type of test report on this line. The mathematical details of each test were described earlier in the Technical Details section of this chapter.

### Lower and Upper Equivalence Limit

These are the limits on bioequivalence in percentage form. As long as the percentage of the treatment formulation of the reference formula is between these limits, the treatment formulation is bioequivalent. These values were set by you. They are not calculated from the data.

### Lower and Upper Confidence Limits

These are the confidence limits on the ratio of mean responses to the two formulations computed from the data. Note that the confidence coefficient is $(1-2\alpha)\times 100\%$. If both of these limits are inside the two equivalence limits, the treatment formulation is bioequivalent to the reference formulation. Otherwise, it is not.

### Equivalent at the 5.0% Sign. Level?

This column indicates whether bioequivalence can be concluded.

# Equivalence Based on Schuirmann's Two One-Sided Hypothesis Tests

| Test Type | Lower Test T Value | Upper Test T Value | 5.0% Cutoff T Value | DF | Equivalent at the 5.0% Sign. Level? |
|---|---|---|---|---|---|
| Schuirmann's 2 1-Sided Tests | 3.81 | -5.04 | 1.72 | 22 | Yes |

**Interpretation of the Above Report**
Average bioequivalence of the two treatments was found at the 0.0500 significance level using Schuirmann's two one-sided t-tests procedure. The probability level of the t-test of whether the treatment mean is not too much lower than the reference mean is 0.0005. The probability level of the t-test of whether the treatment mean is not too much higher than the reference mean is 0.0000. Since both of these values are less than 0.0500, the null hypothesis of average bioinequivalence was rejected in favor of the alternative hypothesis of average bioequivalence. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

This report provides the results of Schuirmann's two one-sided hypothesis tests procedure.

### Test Type

This is the type of test reported on this line. The mathematical details of this test were described earlier in the Technical Details section of this chapter.

### Lower and Upper Test T Value

These are the values of $T_L$ and $T_U$, the two one-sided test statistics.

### 5% Cutoff T Value

This is the *T* value that marks significance or non-significance. If the absolute values of both $T_L$ and $T_U$ are greater than this value, the treatment formulation is bioequivalent. Otherwise, it is not. This *T* value is based on the degrees of freedom and on $\alpha$.

### DF

This is the value of the degrees of freedom. In this case, the value of the degrees of freedom is $n_1 + n_2 - 2$.

### Equivalent at the 5.0% Sign. Level?

This column indicates whether bioequivalence is concluded.

## Equivalence Based on Two One-Sided Wilcoxon-Mann-Whitney Tests

| Test Type | Lower Sum Ranks | Lower Prob Level | Upper Sum Ranks | Upper Prob Level | Equivalent at the 5.0% Sign. Level? |
|---|---|---|---|---|---|
| 2 1-Sided MW Tests | 207.00 | 0.0002 | 91.00 | 0.0001 | Yes |

**Interpretation of the Above Report**
Average bioequivalence of the two treatments was found at the 0.0500 significance level using the nonparametric version of Schuirmann's two one-sided tests procedure which is based on the Wilcoxon-Mann-Whitney test. The probability level of the test of whether the treatment mean is not too much lower than the reference mean is 0.0002. The probability level of the test of whether the treatment mean is not too much higher than the reference mean is 0.0001. Since both of these values are less than 0.0500, the null hypothesis of average bioinequivalence was rejected in favor of the alternative hypothesis of average bioequivalence. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

This report provides the results of the nonparametric version of Schuirmann's two one-sided hypothesis tests procedure.

### Test Type

This is the type of test reported on this line. The mathematical details of this test were described earlier in the Technical Details section of this chapter.

### Lower and Upper Sum Ranks

These are sum of the ranks for the lower and upper Mann-Whitney tests.

### Lower and Upper Prob Level

These are the upper and lower significance levels of the two one-sided Wilcoxon-Mann-Whitney tests. Bioequivalence is indicated when both of these values are less than a given level of $\alpha$.

### Equivalent at the 5.0% Sign. Level?

This column indicates whether bioequivalence is concluded.

## Equivalence Based on Anderson and Hauck's Hypothesis Test

| Test Type | Pr(-TL) | Pr(TU) | Prob Level | Equivalent at the 5.0% Sign. Level? |
|---|---|---|---|---|
| Anderson and Hauck's Test | 0.0005 | 0.0000 | 0.0005 | Yes |

**Interpretation of the Above Report**
Average bioequivalence of the two treatments was found at the 0.0500 significance level using Anderson and Hauck's test procedure. The actual probability level of the test was 0.0005. This experiment used a 2x2 cross-over design with 12 subjects in sequence 1 and 12 subjects in sequence 2.

This report provides the results of Anderson and Hauck's hypothesis test procedure.

### Test Type

This is the type of test reported on this line. The mathematical details of this test were described earlier in the Technical Details section of this chapter.

### Pr(-TL) and Pr(TU)

These values are subtracted to obtain the significance level of the test.

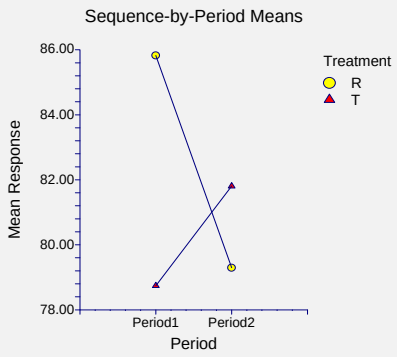
### Prob Level

This is the significance level of the test. Bioequivalence is indicated when this value is less than a given level of $\alpha$.

### Equivalent at the 5.0% Sign. Level?

This column indicates whether bioequivalence is concluded.

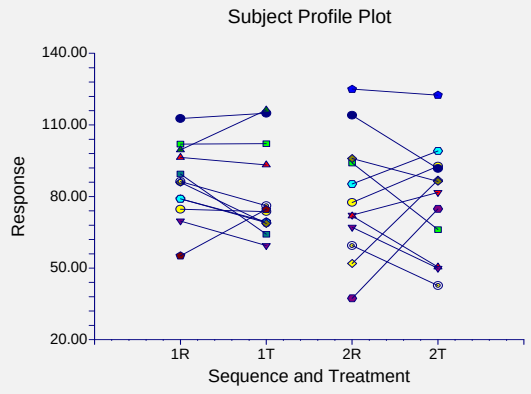
## Plot of Sequence-by-Period Means




The sequence-by-period means plot shows the mean responses on the vertical axis and the periods on the horizontal axis. The lines connect like treatments. The distance between these lines represents the magnitude of the treatment effect.

If there is no period, carryover, or interaction effects, two horizontal lines will be displayed. The tendency for both lines to slope up or down represents period and carryover effects. The tendency for the lines to cross represents period-by-treatment interaction. This is also a type of carryover effect.

## Plot of Subject Profiles

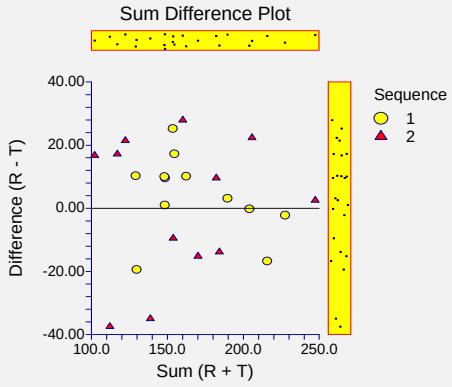


The profile plot displays the raw data for each subject. The response variable is shown along the vertical axis. The two sequences are shown along the horizontal axis. The data for each subject is depicted by two points connected by a line. The subject's response to the reference formulation is shown first followed by their response to the treatment formulation. Hence, for sequence 2, the results for the first period are shown on the right and for the second period on the left.

This plot is used to develop a feel for your data. You should view it first as a tool to check for outliers (points and subjects that are very different from the majority). Note that outliers should be removed from the analysis only if a reason can be found for their deletion. Of course, the first step in dealing with outliers is to double-check the data values to determine if a typing error might have caused them. Also, look for subjects whose lines exhibit a very different pattern from the rest of the subjects in that sequence. These might be a signal of some type of data-recording or data-entry error.

The profile plot allows you to assess the consistency of the responses to the two treatments across subjects. You may also be able to evaluate the degree to which the variation is equal in the two sequences.

## Plot of Sums and Differences



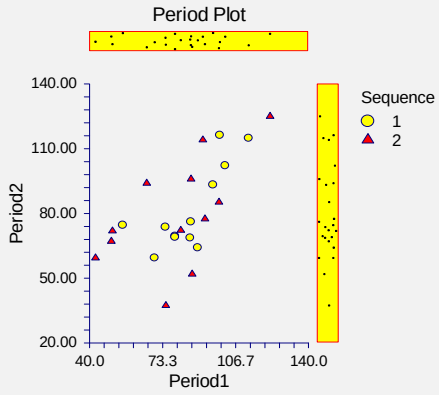

The sums and differences plot shows the sum of each subject's two responses on the horizontal axis and the difference between each subject's two responses on the vertical axis. Dot plots of the sums and differences have been added above and to the right, respectively.

Each point represents the sum and difference of a single subject. Different plotting symbols are used to denote the subject's sequence. A horizontal line has been added at zero to provide an easy reference from which to determine if a difference is positive (favors treatment R) or negative (favors treatment T).

The degree to which the plotting symbols tend to separate along the horizontal axis represents the size of the carryover effect. The degree to which the plotting symbols tend to separate along the vertical axis represents the size of the treatment effect.

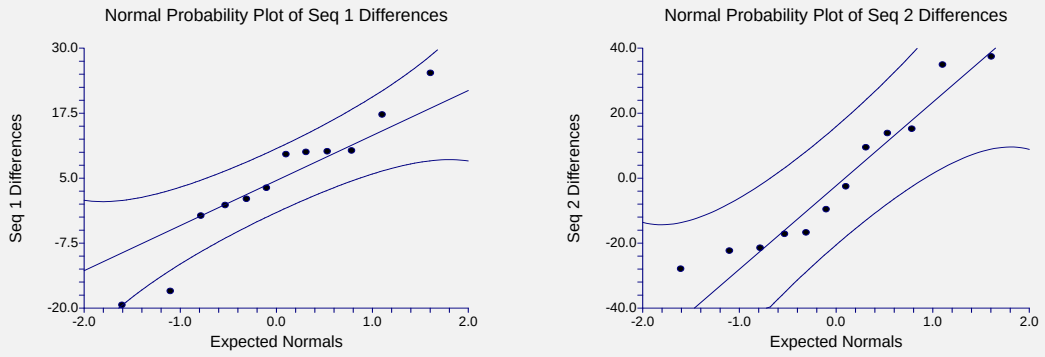
Outliers are easily detected on this plot. Outlying subjects should be reviewed for data-entry errors and for special conditions that might have caused their responses to be unusual. Outliers should not be removed from an analysis just because they are different. A compelling reason should be found for their removal and the removal should be well documented.

## Period Plot




The Period Plot displays a subject's period 1 response on the horizontal axis and their period 2 response on the vertical axis. The plotting symbol is the sequence number. The plot is used to find outliers and other anomalies.

## Probability Plots




These plots show the differences (*P*1-*P*2) on the vertical axis and values on the horizontal axis that would be expected if the differences were normally distributed. The first plot shows the differences for sequence 1 and the second plot shows the differences for sequence 2.

If the assumption of normality holds, the points should fall along a straight line. The degree to which the points are off the line represents the degree to which the normality assumption does not hold. Since the normality of these differences is assumed by the *t*-test used to test for a difference between the treatments, these plots are useful in assessing whether that assumption is valid.

If the plots show a pronounced pattern of non-normality, you might try taking the square roots or the logs of the responses before beginning the analysis.

# Chapter 240

# Nondetects Analysis

## Introduction

This procedure computes summary statistics, generates EDF plots, and computes hypothesis tests appropriate for two or more groups for data with nondetects (left-censored) values. Following the recommendation of Helsel (2005), pp. 77-78, the methods for this procedure are valid only if fewer than 50% of the values are nondetects (left-censored).

Nondetects analysis is the analysis of data in which one or more of the values cannot be measured exactly because they fall below one or more detection limits. Detection limits often arise in environmental studies because of the inability of instruments to measure small concentrations. Some examples of sampling scenarios that lead to datasets with nondetects values are finding pesticide concentrations in water, determining chemical composition of soils, or establishing the number of particulates of a compound in the air.

A common practice for dealing with values which fall below the detection threshold is substitution. Often, each value which is below the detection limit is substituted with one half the detection limit. Summary statistics and comparisons are then carried out using standard techniques (means, confidence intervals, t-tests, ANOVA, etc.) with the substituted data. Helsel (2005) warns of the potential data analysis biases that result if nondetects values are substituted. He particularly warns about the arbitrariness of substituting one half the detection limit (or zero, or the detection limit). Alternatively, techniques based on survival analysis methods have been developed for appropriate use of the information contained in the nondetected observations. The general approach is to convert the nondetects data (left-censored) to survival data (right-censored), use the survival analysis techniques on the newly created survival data, and then convert the survival summaries back to original scale (In *NCSS*, these conversions are performed automatically). The resulting summary statistics and hypothesis tests are analogs to the common techniques, but which appropriately account for nondetected observations. For example, medians are used rather than means, EDF plots replace box plots and histograms, and logrank tests are used instead of two-sample t-tests and ANOVA.

The technical details of survival analysis are found in the Kaplan-Meier Survival Curves chapter. For a complete account of nondetects analysis, we suggest the book by Helsel (2005).

# Technical Details

## Flipping Constant

To convert nondetects data to the format of survival data, each response, including nondetected values, must be subtracted from a suitable flipping constant. The flipping constant can be any number which is larger than the maximum of the nondetects data. The resulting right-censored data are

$$Flip_i = M - x_i,$$

where $M$ is the flipping constant and the $x_i$ are the original observations.

For example, consider the first 10 of 25 dioxin concentrations (fg/cubic meter) with lower detection limit 50 fg/cubic meter (these data can be found in the DIOXIN dataset):

**DIOXIN dataset (subset)**

| Dioxin |
|---|
| 391 |
| 724 |
| 603 |
| 50 |
| 482 |
| 656 |
| 50 |
| 797 |
| 190 |
| 444 |

A suitable flipping constant is any value larger then the maximum value. Suppose $M$ = 1000 is arbitrarily chosen as the flipping constant. The flipped data would then become

| Dioxin | *M* – Dioxin | Flip |
|---|---|---|
| 391 | 1000 – 391 | 609 |
| 724 | 1000 – 724 | 276 |
| 603 | 1000 – 603 | 397 |
| <50 | 1000 – <50 | >950 |
| 482 | 1000 – 482 | 518 |
| 656 | 1000 – 656 | 344 |
| <50 | 1000 – <50 | >950 |
| 797 | 1000 – 797 | 203 |
| 190 | 1000 – 190 | 810 |
| 444 | 1000 – 444 | 556 |
| . | . | . |
| . | . | . |
| . | . | . |

The flipped data is now in the survival data format.

Once the data are converted to the survival data format, the nonparametric Kaplan-Meier methods can be used for estimating summary statistics (i.e., median, quantiles, standard errors, confidence limits), and for group comparisons. The summary statistics of location (i.e., median, quantiles, and confidence limits) are converted back to the original scale using the same flipping constant *M*. For example, to convert the median of the survival data to the median of the original units, the formula

$$Median = M - SurvivalMedian$$

is used. For the Dioxin data, the survival median (of the flipped data) is 556 fg/cubic meter. The median on the original scale would then be Median = 1000 – 556 = 444 fg/cubic meter. The standard error statistics for the flipped survival data are the same as those of the original scale, and need not be converted. All of the calculations involving conversion and re-conversion based on the flipping constant are done automatically in *NCSS*.

## The Empirical Distribution Function (EDF)

The empirical distribution function (EDF) provides an approximation of the true cumulative distribution function of the measured response. It is useful for viewing or obtaining sample percentiles (quantiles) for each of the observed responses. The EDF is produced using the Kaplan-Meier product-limit estimator (estimated survival distribution) of the flipped data. The resulting survival distribution is then converted to the EDF by re-subtracting all values from the flipping constant. We now examine the technical details of the estimation of the survival distribution.

## Hypothesis Tests

This section presents methods for testing that the distribution functions of two or more populations are equal. The null hypothesis is that the distribution functions of all populations are equal at all values greater than the minimum observed value. The alternative hypothesis is that at least two of the distribution functions are different at some value greater than the observed minimum value.

Five different choices of tests are available in ***NCSS*** to test the above hypotheses. The tests differ in the manner in which different responses are weighted. The most commonly used test is the logrank test, which has equal weighting. The other four tests shift the heaviest weighting to the larger or smaller responses. Although five tests are displayed, only one should be used. Because of the different weighting patterns, they will often give quite different results. The test that will be used should be justified and designated before viewing the data or test results.

The following table describes the weighting scheme for each of these tests.

| Test | Comments |
|---|---|
| Logrank | This is the most commonly used test and the one we recommend. Equal weights across all times are used. |
| Gehan | Places very heavy weight on large responses. |
| Tarone-Ware | Places heavy weight on small responses. |
| Peto-Peto | Places a little more weight on large responses. |
| Modified Peto-Peto | Places a little more weight on large responses. |

## Data Structure

Nondetects datasets are specified using up to four components: the response value (e.g., concentration or amount), an optional indicator of whether or not each observation was detected, an optional group specification, and an optional frequency (count) specification. If no detection indicator is included, all response values represent detected responses. If there is no group specification, a single group is assumed. If the frequency (count) variable is omitted, all counts are assumed to be one.

## Sample Dataset

The table below shows a dataset (fictitious) reporting sediment arsenic concentrations for three different regions of a lake. A single sample was taken from each of twenty randomly selected locations of each region. In this dataset, the response is the concentration of arsenic in mg/Kg (dry weight). The instruments used in the study to determine arsenic concentration are unable to detect concentrations below 10 mg/Kg. A value of zero in the ANondet column indicates arsenic was detected. A value of one in the ANondet column indicates arsenic was not detected. These data are contained in the ARSENIC dataset.

**ARSENIC dataset (subset)**

| Arsenic | ANondet | Region |
|---|---|---|
| 14 | 0 | 1 |
| 10 | 1 | 1 |
| 31 | 0 | 1 |
| 26 | 0 | 1 |
| 10 | 1 | 1 |
| . | . | . |
| . | . | . |
| . | . | . |
| 15 | 0 | 2 |
| 10 | 1 | 2 |
| 25 | 0 | 2 |
| 21 | 0 | 2 |
| 27 | 0 | 2 |
| . | . | . |
| . | . | . |
| . | . | . |
| 29 | 0 | 2 |
| 26 | 0 | 2 |
| 18 | 0 | 3 |
| 26 | 0 | 3 |
| . | . | . |
| . | . | . |
| . | . | . |

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

This panel specifies the variables used in the analysis.

### Response Variable

#### Response Variable

The values of this variable represent either the magnitude of a detected observations or detection limits, depending on the corresponding values of the Nondetection (Censor) Variable.

The values in this variable must be greater than zero. If the value is missing or non-positive, it is not used during the estimation phase.

### Nondetection Variable

#### Nondetection (Censor) Variable

The values in this variable indicate whether the value of the Response Variable represents a nondetected (censored) observation or a detected observation. When a particular value of this variable indicates a Nondetect, the corresponding value of the Response Variable represents a lower detection limit.

These values may be text or numeric. The interpretation of these codes is specified by the 'Detected' and 'Not Detected' (Censored) options to the right of this option.

Only two values are used, the Detected value and the Not Detected value. The Unknown Censor option specifies what is to be done with values that do not match either the Detected value or the Not Detected value.

Rows with missing values (blanks) in this variable are omitted from the estimation phase, but results are shown in any reports that output predicted values.

#### Detected

When this value is encountered under the Nondetection (Censor) Variable it indicates that the value under the Response Variable was observed or detected. The value may be a number or a letter.

We suggest the letter 'D' or the number '0' when you are in doubt as to what to use.

A detected observation is one in which the value was measured exactly; for example, the concentration was such that the instrument was able to measure it.

#### Not Detected

When this value is encountered under the Nondetection (Censor) Variable it indicates that the value under the Response Variable was not actually observed (i.e., a nondetect) but represents a lower detection limit. That is, the observation is left-censored, and the actual value of the response is something below the detection limit.

The value may be a number or a letter. We suggest the letter 'N' or the number '1' when you are in doubt as to what to use.

A nondetect is a response in which the value was not measured exactly; for example, the concentration was such that the instrument was not able to measure it.

### Unknown Censor

This option specifies what the program is to assume about observations whose Nondetection (Censor) Variable value is not equal to either the Detected code or the Not Detected code. Note that observations with missing Nondetection (Censor) values are always treated as missing.

- **Not Detected**

  Observations with unknown Nondetection (Censor) Variable values are assumed to be nondetects (censored).

- **Detected**

  Observations with unknown Nondetection (Censor) Variable values are assumed to be detected.

- **Missing**

  Observations with unknown Nondetection (Censor) Variable values are assumed to be missing and those rows are omitted from the analysis.

## Frequency Variable

### Frequency Variable

Specify an optional variable containing the number of observations (cases) represented by each row.

If this variable is left blank, each row of the database is assumed to represent one observation.

## Group Variable

### Group Variable

An optional categorical (grouping) variable may be specified. If it is used, a separate analysis is conducted for each unique value of this variable. A variable must be entered here to generate log rank test comparisons.

## Options

### Alpha Level

This is the value to which probability levels are compared for testing hypotheses. Also, one minus alpha is the confidence level used for confidence intervals. For example, if you specify 0.04 here, then 96% confidence limits will be calculated.

A value of .05 is historically the most commonly used. For hypothesis testing, this value represents a 1 in 20 chance of falsely rejecting the null hypothesis. For confidence intervals, this corresponds to a chance of 1 out of 20 of creating an interval that does not contain the true parameter. Now, values other than 0.05 are often recommended or required by journals or institutions. Typical values range from 0.001 to 0.20.

### Confidence Limits

This option specifies the method used to estimate the confidence limits. The options are:

- **Linear**

  This is the classical method, which uses Greenwood's estimate of the variance.

- **Log Transform**

  This method uses the logarithmic transformation of Greenwood's variance estimate. It produces better limits than the Linear method and has better small sample properties.

- **ArcSine**

  This method uses the arcsine square-root transformation of Greenwood's variance estimate to produce better limits.

## Reports Tab

The following options control which reports and plots are displayed.

### Select Reports

#### Data Summary Section ... Logrank Test Detail

Specify whether to display the indicated reports.

#### Specific Responses

Specify a list of values for which cumulative proportions are to be calculated. These values are used only if the 'Specific Response Detail' box is checked.

Numbers are separated by blanks or commas. Specify sequences with a colon, putting the increment inside parentheses. For example: 5:25(5) means 5 10 15 20 25.

Use '(10)' alone to specify ten, equal-spaced values between zero and the maximum.

Only positive values may be entered here.

#### Quantiles

Specify a list of quantiles (percentiles) for which the estimated response is to be calculated. These values are used only if the 'Quantiles of Responses' box is checked.

Numbers are separated by blanks or commas in this list. Specify sequences with a colon, putting the increment inside parentheses. For example: 5:25(5) means 5 10 15 20 25 and 1:5(2),10:20(2) means 1 3 5 10 12 14 16 18 20.

All values in the list must be between 0 and 100.

### Select Plots

#### EDF Plot

Specify whether to display the indicated plot.

## Select Plots – Plots Displayed

### Individual-Group Plots

When checked, this option specifies that a separate chart of each designated type is displayed.

### Combined Plot

When checked, this option specifies that a chart combining all groups is to be displayed.

## Report Options

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. Single precision is for reporting purposes only.

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Value Labels

This option lets you select whether to display only values, only value labels, or both for values of the group variable. Use this option if you want to automatically attach labels to the values of the group variable (such as 1=Male, 2=Female, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

## Report Options – Decimal Places

### Response ... Chi-Square Decimals

This option specifies the number of decimal places shown on reported values.

## Plot Options – Plot Arrangement

### Two Plots Per Line

When unchecked, one large plot is displayed per line. When checked, two smaller plots are displayed per line.

# EDF Plots Tab

The following options control the EDF plots that are displayed.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum and Maximum

These options specify the minimum and maximum values to be displayed on the vertical (Y) and horizontal (X) axis. If left blank, these values are calculated from the data.

### Tick Label Settings...

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Ticks: Major and Minor

These options set the number of major and minor tick marks displayed on each axis.

### Show Grid Lines

These check boxes indicate whether the grid lines should be displayed.

## Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Censor Tickmarks

This option indicates the size of the tickmarks (if any) showing where the nondetected (censored) points fall on the EDF curve. The values are at a scale of 1000 equals one inch.

We recommend that you use ‘0’ to indicate no marks or ‘100’ to display the marks.

## Plot Settings – Plot Contents

These options control objects that are displayed on all plots.

### Function Line

Indicate whether to display the EDF curve on the plots.

### C.L. Lines

Indicate whether to display the confidence limits of the estimated function on the plots.

### Legend

Specifies whether to display the legend.

### Legend Text

Specifies legend label. If *{G}* is entered here, *{G}* is replaced by the name of the group variable.

## Titles

### Title Line 1 and 2

These are the text lines of the titles. The characters *{X}*, *{G}*, and *{Z}* are replaced by appropriate names. The color or font of the text may be specified here by pressing the button to the right of the field.

# Lines Tab

These options specify the attributes of the lines used for each group in the EDF plots.

## Plotting Lines

### Line 1 - 15

These options specify the color, width, and pattern of the lines used in the plots of each group. The first line is used by the first group, the second line by the second group, and so on. These line attributes are provided to allow the various groups to be indicated on black-and-white printers.

Clicking on a line box (or the small button to the right of the line box) will bring up a window that allows the color, width, and pattern of the line to be changed.

# Storage Tab

These options let you specify if, and where on the database, various statistics are stored.

*Warning: If statistics are stored into columns which already contain data, any data in these columns is replaced by the new statistics data. Be careful not to specify variables that contain important data.*

## Data Storage Options

### Storage Option

This option controls whether the values indicated below are stored on the database when the procedure is run.

- **Do not store data**
  No data are stored even if they are checked.

- **Store in empty columns only**
  The values are stored in empty columns only. Columns containing data are not used for data storage, so no data can be lost.

- **Store in designated columns**
  Beginning at the *First Storage Variable*, the values are stored in this column and those to the right. If a column contains data, the data are replaced by the storage values. Care must be used with this option because it cannot be undone.

### Store First Variable In

The first item is stored in this variable. Each additional item that is checked is stored in the variables immediately to the right of this variable.

Leave this value blank if you want the data storage to begin in the first blank column on the right-hand side of the data.

Warning: Any existing data in these variables is automatically replaced.

### Data Storage Options – Select Items to Store

#### Response Group ... UCL of P(R)

Indicate whether to store these values, beginning at the variable indicated by the *Store First Variable In* option.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

#### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

## Example 1 – Analysis of Data with Nondetects

This section presents an example of how to analyze a typical set of nondetects data. Twenty-five air quality locations were randomly chosen to determine dioxin concentration (fg/cubic meter). The lower detection limit of the measurement instrument is 50 fg/cubic meter. Four of the 25 concentrations were not detected, and thus, are known only to be less than 50.

The data used are recorded in the DIOXIN dataset.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Nondetects Analysis window.

**1 Open the DIOXIN dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **DIOXIN.S0**.
- Click **Open**.

**2 Open the Nondetects Analysis window.**

- On the menus, select **Analysis**, then **Nondetects**, then **Nondetects Analysis**. The Nondetects Analysis procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Nondetects Analysis window, select the **Variables tab**.
- Set the **Response Variable** to **Dioxin**.
- Set the **Nondetection (Censor) Variable** to **DNondet**.
- Set **Detected** to **0**.
- Set **Not Detected** to **1**.

**4 Specify the reports.**

- Select the **Reports tab**.
- Set the **Specific Responses** box to **100:500(100)**.

**5 Adjust the plots.**

- Select the **EDF Plots tab**.
- Under Vertical Axis, click on **Tick Label Settings**.
- Change **Decimals** to **2**.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Data Summary Section

**Data Summary Section**

| Type | Rows | Count | Minimum | Maximum |
|---|---|---|---|---|
| Detected | 21 | 21 | 94 | 801 |
| Not Detected | 4 | 4 | 50 | 50 |
| Total | 25 | 25 | 50 | 801 |

**Data Summary Section: Response Quartiles**

| Quartile | Estimate | Lower 95.0% C.L. | Upper 95.0% C.L. |
|---|---|---|---|
| First (Q1) | 190.000 | 50.000 | 438.000 |
| Median (Q2) | 444.000 | 199.000 | 603.000 |
| Third (Q3) | 603.000 | 455.000 | 724.000 |

This report displays a summary of the amount of data that were analyzed and the three quartiles. Scan this report to determine if there were any obvious data errors by double checking the counts and the minimum and maximum responses.

## Specific Response Detail: Estimated Cumulative Proportion

| Response (R) | Cumulative Proportion P(R) | Standard Error of P(R) | Lower 95.0% C.L. for P(R) | Upper 95.0% C.L. for P(R) | Cum. Count |
|---|---|---|---|---|---|
| 100.000 | 0.2000 | 0.0800 | 0.0432 | 0.3568 | 5 |
| 200.000 | 0.3200 | 0.0933 | 0.1371 | 0.5029 | 8 |
| 300.000 | 0.3200 | 0.0933 | 0.1371 | 0.5029 | 8 |
| 400.000 | 0.4400 | 0.0993 | 0.2454 | 0.6346 | 11 |
| 500.000 | 0.6000 | 0.0980 | 0.4080 | 0.7920 | 15 |

This report displays the Kaplan-Meier cumulative proportions at the specified responses. The standard error and confidence limits are also shown.

### Response (R)

This is the specific response being reported on this line. The response values were specified in the Specific Responses box under the Reports tab.

### Cumulative Proportion P(R)

This is the estimated proportion of responses less than the specified response (R).

### Standard Error of P(R)

This is the estimated standard error, the square root of the variance estimate given by Greenwood's formula.

### Lower and Upper Confidence Limits for S(T)

The lower and upper confidence limits provide a pointwise confidence interval for the cumulative proportion at each response. These limits are constructed so that the probability that the true proportion lies between them is $1-\alpha$.

Three difference confidence intervals are available. All three confidence intervals perform similarly for large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits are given in the Kaplan-Meier Survival Curves chapter and are not repeated here.

### Cumulative Count

This value is the number of less than or equal to the specified response (R).

## Quantiles of Responses

| Proportion of Response | Estimated Quantile | Lower 95.0% C.L. Quantile | Upper 95.0% C.L. Quantile |
|---|---|---|---|
| 0.0500 | | 50.000 | 126.000 |
| 0.1000 | | 50.000 | 190.000 |
| 0.1500 | | 50.000 | 329.000 |
| 0.2000 | 126.000 | 50.000 | 336.000 |
| 0.2500 | 190.000 | 50.000 | 438.000 |
| 0.3000 | 199.000 | 50.000 | 444.000 |
| 0.3500 | 329.000 | 94.000 | 455.000 |
| 0.4000 | 391.000 | 126.000 | 482.000 |
| 0.4500 | 438.000 | 190.000 | 537.000 |
| 0.5000 | 444.000 | 199.000 | 603.000 |
| 0.5500 | 455.000 | 336.000 | 603.000 |
| 0.6000 | 537.000 | 391.000 | 626.000 |
| 0.6500 | 557.000 | 438.000 | 656.000 |
| 0.7000 | 603.000 | 444.000 | 724.000 |
| 0.7500 | 603.000 | 455.000 | 724.000 |
| 0.8000 | 656.000 | 537.000 | 764.000 |
| 0.8500 | 724.000 | 557.000 | 797.000 |
| 0.9000 | 764.000 | 603.000 | 801.000 |
| 0.9500 | 797.000 | 626.000 | 801.000 |

This report displays the estimated quantiles for various response proportions. For example, it gives the median response if it can be estimated.

### Proportion of Response

This is the response proportion that is reported on this line. The proportion values were specified in the Quantiles box under the Reports tab.

### Estimated Quantile

This is the response value corresponding to the response proportion. For example, this table estimates that 65% of the concentrations are less than or equal to 557 fg/m$^3$.

### Lower and Upper Confidence Limits on Quantiles

These values provide a pointwise $100(1-\alpha)\%$ confidence interval for the estimated quantiles. For example, if the proportion of response 0.50, this provides a confidence interval for the median survival time.

Three methods are available for calculating these confidence limits. The method is designated under the Variables tab in the Confidence Limits box. The formulas for these confidence limits are given in the Kaplan-Meier Survival Curves chapter and are not repeated here.

Because of censoring, estimates and confidence limits are not available for all response proportions.

## Response Detail

| Response (R) | Cumulative Proportion P(R) | Standard Error of P(R) | Lower 95.0% C.L. for P(R) | Upper 95.0% C.L. for P(R) | Cum. Count | Count |
|---|---|---|---|---|---|---|
| <50.000 | | | | | 4 | 4 |
| 94.000 | 0.1600 | 0.0733 | 0.0163 | 0.3037 | 5 | 1 |
| 126.000 | 0.2000 | 0.0800 | 0.0432 | 0.3568 | 6 | 1 |
| 190.000 | 0.2400 | 0.0854 | 0.0726 | 0.4074 | 7 | 1 |
| 199.000 | 0.2800 | 0.0898 | 0.1040 | 0.4560 | 8 | 1 |
| 329.000 | 0.3200 | 0.0933 | 0.1371 | 0.5029 | 9 | 1 |
| 336.000 | 0.3600 | 0.0960 | 0.1718 | 0.5482 | 10 | 1 |
| 391.000 | 0.4000 | 0.0980 | 0.2080 | 0.5920 | 11 | 1 |
| 438.000 | 0.4400 | 0.0993 | 0.2454 | 0.6346 | 12 | 1 |
| 444.000 | 0.4800 | 0.0999 | 0.2842 | 0.6758 | 13 | 1 |
| 455.000 | 0.5200 | 0.0999 | 0.3242 | 0.7158 | 14 | 1 |
| 482.000 | 0.5600 | 0.0993 | 0.3654 | 0.7546 | 15 | 1 |
| 537.000 | 0.6000 | 0.0980 | 0.4080 | 0.7920 | 16 | 1 |
| 557.000 | 0.6400 | 0.0960 | 0.4518 | 0.8282 | 17 | 1 |
| 603.000 | 0.6800 | 0.0933 | 0.4971 | 0.8629 | 19 | 2 |
| 626.000 | 0.7600 | 0.0854 | 0.5926 | 0.9274 | 20 | 1 |
| 656.000 | 0.8000 | 0.0800 | 0.6432 | 0.9568 | 21 | 1 |
| 724.000 | 0.8400 | 0.0733 | 0.6963 | 0.9837 | 22 | 1 |
| 764.000 | 0.8800 | 0.0650 | 0.7526 | 1.0000 | 23 | 1 |
| 797.000 | 0.9200 | 0.0543 | 0.8137 | 1.0000 | 24 | 1 |
| 801.000 | 0.9600 | 0.0392 | 0.8832 | 1.0000 | 25 | 1 |

This report displays the Kaplan-Meier product-limit distribution values along with confidence limits. The formulas used are given in the Kaplan-Meier Survival Curves chapter.

### Response (R)

This is the response being reported on this line. The response are the unique responses that occurred in the data.

Note that observations which are nondetects are marked with a less than sign (<). Estimated proportions are not calculated for nondetects observations.

### Cumulative Proportion P(R)

This is the estimated proportion of responses less than the response (R).

### Standard Error of S(T)

This is the estimated standard error, the square root of the variance estimate given by Greenwood's formula.

### Lower and Upper Confidence Limits for S(T)

The lower and upper confidence limits provide a pointwise confidence interval for the cumulative proportion at each response. These limits are constructed so that the probability that the true proportion lies between them is $1-\alpha$.

Three difference confidence intervals are available. All three confidence intervals perform similarly for large samples. The linear (Greenwood) interval is the most commonly used. However, the log-transformed and the arcsine-square intervals behave better in small to moderate samples, so they are recommended. The formulas for these limits are given in the Kaplan-Meier Survival Curves chapter and are not repeated here.

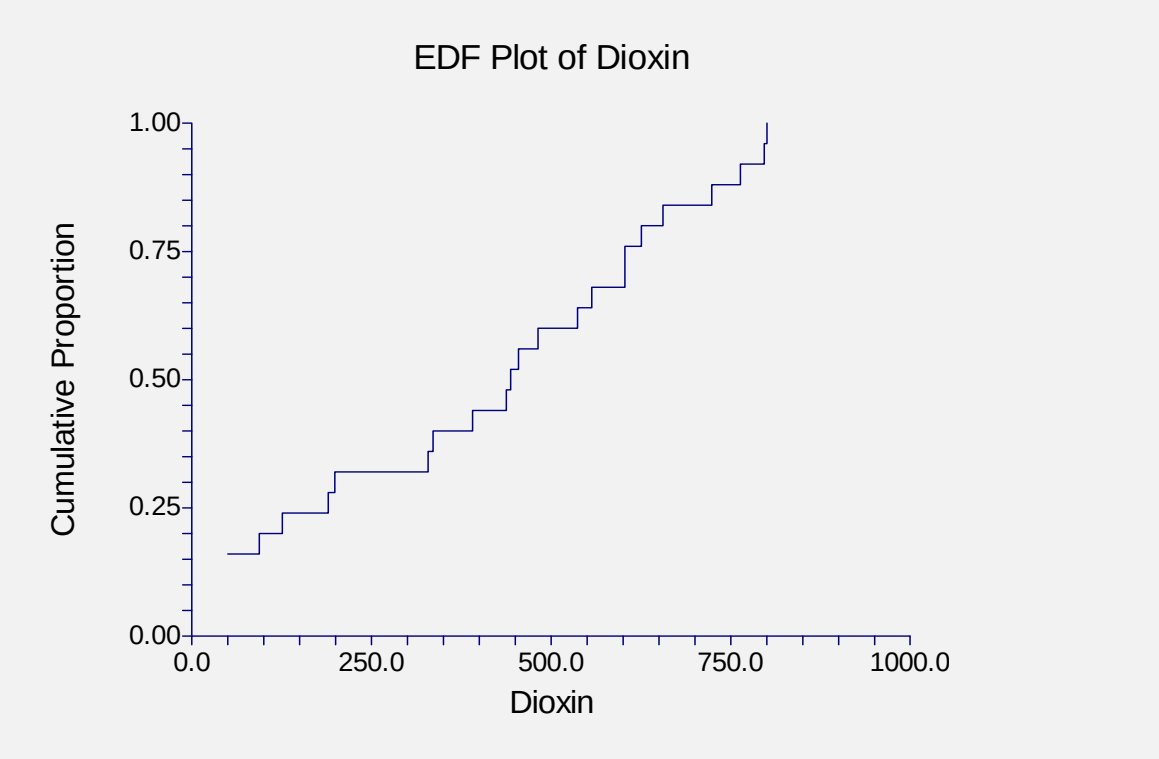
### Cumulative Count

This value is the number of less than or equal to the specified response (R).

### Count

This is the number of observations with this specific response value.

## EDF Plot




This plot shows the empirical distribution function (EDF). If there are several groups, a separate line is drawn for each group.

# Example 2 – Group Comparisons with Nondetects

The research purpose of this example is comparing sediment arsenic concentrations for three different regions of a lake. A single sample was taken from each of twenty randomly selected locations of each region. The response is the concentration of arsenic in mg/Kg (dry weight). The instruments used in the study to determine arsenic concentration are unable to detect concentrations below 10 mg/Kg.

The data used are recorded in the variables Arsenic, ANondet, and Region of the ARSENIC dataset.

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Nondetects Analysis window.

**1 Open the ARSENIC dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **ARSENIC.S0**.
- Click **Open**.

**2 Open the Nondetects Analysis window.**

- On the menus, select **Analysis**, then **Nondetects**, then **Nondetects Analysis**. The Nondetects Analysis procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Nondetects Analysis window, select the **Variables tab**.
- Set the **Response Variable** to **Arsenic**.
- Set the **Nondetection (Censor) Variable** to **ANondet**.
- Set the **Group Variable** to **Region**.

**4 Specify the reports.**

- On the Nondetects Analysis window, select the **Reports tab**.
- Check the **Logrank Test Summary** box.
- Check the **Logrank Test Detail** box.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

# Logrank Tests Section

**Hypotheses**
H0: Distribution Functions are Equal Among Groups
HA: At Least One Group Distribution Functions Differs

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha = 0.05) |
|---|---|---|---|---|
| Logrank | 26.680 | 2 | 0.0000 | Yes |
| Gehan-Wilcoxon | 35.265 | 2 | 0.0000 | Yes |
| Tarone-Ware | 32.241 | 2 | 0.0000 | Yes |
| Peto-Peto | 35.479 | 2 | 0.0000 | Yes |
| Mod. Peto-Peto | 35.589 | 2 | 0.0000 | Yes |

**Multiple Pairwise Tests Section**
**Hypotheses**
H0: Distribution Functions are Equal
HA: Distribution Functions Differ

**Group Pair Tested: 1 vs. 2**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| Logrank | 0.374 | 1 | 0.5409 | No | 1.0000 | No |
| Gehan-Wilcoxon | 0.326 | 1 | 0.5683 | No | 1.0000 | No |
| Tarone-Ware | 0.389 | 1 | 0.5327 | No | 1.0000 | No |
| Peto-Peto | 0.267 | 1 | 0.6055 | No | 1.0000 | No |
| Mod. Peto-Peto | 0.265 | 1 | 0.6069 | No | 1.0000 | No |

**Group Pair Tested: 1 vs. 3**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| Logrank | 16.239 | 1 | 0.0001 | Yes | 0.0002 | Yes |
| Gehan-Wilcoxon | 19.657 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Tarone-Ware | 18.787 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Peto-Peto | 19.418 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Mod. Peto-Peto | 19.457 | 1 | 0.0000 | Yes | 0.0000 | Yes |

**Group Pair Tested: 2 vs. 3**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| Logrank | 15.978 | 1 | 0.0001 | Yes | 0.0002 | Yes |
| Gehan-Wilcoxon | 20.474 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Tarone-Ware | 19.109 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Peto-Peto | 20.391 | 1 | 0.0000 | Yes | 0.0000 | Yes |
| Mod. Peto-Peto | 20.453 | 1 | 0.0000 | Yes | 0.0000 | Yes |

Notes:
The most commonly used test is the Logrank test.

This report gives the results of the five logrank type tests that are provided by this procedure. We strongly suggest that you select the test that will be used before viewing this report. We recommend Logrank test.

The tests are divided into two groups: overall tests and pairwise tests. The overall tests test for significant differences between groups, but do not indicate which groups are different from each other. The pairwise tests indicate which groups have significantly different distribution functions. Adjusted probability levels should be used to account for multiplicity of tests.

### Chi-Square

This is the chi-square value of the test. Each of these tests is approximately distributed as a chi-square in large samples.

### DF

This is the degrees of freedom of the chi-square distribution associated with each test. It is one less than the number of groups being compared in a particular test.

### Prob Level

This is the significance level of the test. If this value is less than than chosen significance level (often 0.05), the test is significant, indicating evidence of a difference in distribution functions. For pairwise tests the Bonferroni adjusted probability level should be used to account for multiple testing.

### Reject H0

This is an indicator based on the comparison of the probability level to the specified alpha. 'Yes' indicates rejection of the null hypothesis (evidence that the true distribution functions are different). 'No' indicates the null hypothesis should not be rejected (not sufficient evidence that the true distribution functions are different).

### Bonferroni Adjusted Prob Level

When more than two groups are compared, the number of pairwise comparisons is greater than one. Bonferroni adjusted probability levels account for the multiplicity of hypothesis tests. The Bonferroni adjustment to the probability level is made by multiplying the given probability level by the number of tests that are performed (with a ceiling of 1.0). In this example, three pairwise comparisons are made. Thus, each probability level is multiplied by three. Any adjusted probability level greater than one is set to one. The Bonferroni adjusted probability level for the last two longrank tests in this example appears to be only two times the base probability level. This is due to rounding. If more decimal places are specified, it is seen that the adjusted probability levels are three times the base probability levels.

## Logrank Test Detail Section

**Logrank Test Detail Section**

| Group | Z-Value | Standard Error | Standardized Z-Value |
|---|---|---|---|
| 1 | -7.561 | 3.398 | -2.225 |
| 2 | -4.484 | 3.380 | -1.327 |
| 3 | 12.044 | 2.340 | 5.146 |

Probability Level was 0.0000

**Gehan-Wilcoxon Test Detail Section**

| Group | Z-Value | Standard Error | Standardized Z-Value |
|---|---|---|---|
| 1 | -349.000 | 132.199 | -2.640 |
| 2 | -270.000 | 132.219 | -2.042 |
| 3 | 619.000 | 104.394 | 5.929 |

Probability Level was 0.0000

**Tarone-Ware Test Detail Section**

| Group | Z-Value | Standard Error | Standardized Z-Value |
|---|---|---|---|
| 1 | -51.277 | 20.460 | -2.506 |
| 2 | -35.076 | 20.428 | -1.717 |
| 3 | 86.353 | 15.249 | 5.663 |

Probability Level was 0.0000

**Peto-Peto Test Detail Section**

| Group | Z-Value | Standard Error | Standardized Z-Value |
|---|---|---|---|
| 1 | -5.568 | 2.114 | -2.634 |
| 2 | -4.453 | 2.114 | -2.106 |
| 3 | 10.021 | 1.684 | 5.949 |

Probability Level was 0.0000

**Mod. Peto-Peto Test Detail Section**

| Group | Z-Value | Standard Error | Standardized Z-Value |
|---|---|---|---|
| 1 | -5.452 | 2.065 | -2.640 |
| 2 | -4.377 | 2.066 | -2.119 |
| 3 | 9.830 | 1.650 | 5.959 |

Probability Level was 0.0000

This report gives the details of each of the five logrank tests that are provided by this procedure. We strongly suggest that you select the test that will be used before viewing this report. We recommend that you use the Logrank test.

### Group

This is the group reported on this line.

### Z-Value

The details of the z-value are given in the Kaplan-Meier Survival Curves chapter and are not repeated here.

### Standard Error

This is the standard error of the above z-value. It is used to standardize the z-values.

### Standardized Z-Value

The standardized z-value is created by dividing the z-value by its standard error. This provides an index number that will usually very between -3 and 3. Extreme values represent groups that are quite different from the typical group, at least at some response values.

## Example 3 – Validation of Summary Statistics using Helsel (2005)

This section presents validation of nondetects analysis summary statistics. Helsel (2005) presents an example on pages 103-113 involving lead concentrations. These data are contained in the LEAD dataset.

On page 108, Helsel (2005) finds the median to be 1 – 0.984483 = 0.015517. The first and third quartiles are 1 – 0.985714 = 0.014286 and 1 – 0.975472 = 0.024528, respectively. The cumulative proportion for a lead concentration of 0.034 is 0.777778. The (B-C Sign) 95% confidence interval for the median lead concentration is presented on page 112 as (0.014, 0.019).

You may follow along here by making the appropriate entries or load the completed template **Example3** from the Template tab of the Nondetects Analysis window.

**1 Open the LEAD dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **LEAD.S0**.
- Click **Open**.

### 2 Open the Nondetects Analysis window.

- On the menus, select **Analysis**, then **Nondetects**, then **Nondetects Analysis**. The Nondetects Analysis procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the Nondetects Analysis window, select the **Variables tab**.
- Set the **Response Variable** to **Lead**.
- Set the **Nondetection (Censor) Variable** to **LNondet**.

### 4 Specify the reports.

- On the Nondetects Analysis window, select the **Reports tab**.
- Uncheck all reports except the **Data Summary Section** and **Response Detail**.
- Change **Decimal Places - Response** to **6**.

### 5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Output

**Data Summary Section**

| Type | Rows | Count | Minimum | Maximum |
|---|---|---|---|---|
| Detected | 12 | 12 | 1.372549E-02 | 0.2689655 |
| Not Detected | 15 | 15 | 0.02 | 0.02 |
| Total | 27 | 27 | 1.372549E-02 | 0.2689655 |

**Data Summary Section: Response Quartiles**

| Quartile | Estimate | Lower 95.0% C.L. | Upper 95.0% C.L. |
|---|---|---|---|
| First (Q1) | **0.014286** | 0.013725 | 0.018644 |
| Median (Q2) | **0.015517** | **0.014286** | **0.018644** |
| Third (Q3) | **0.024528** | 0.015517 | 0.106061 |

**Response Detail**

| Response (R) | Cumulative Proportion P(R) | Standard Error of P(R) | Lower 95.0% C.L. for P(R) | Upper 95.0% C.L. for P(R) | Cum. Count | Count |
|---|---|---|---|---|---|---|
| 0.013725 | 0.0000 | | | | 1 | 1 |
| 0.014286 | 0.1759 | 0.1539 | 0.0000 | 0.4776 | 2 | 1 |
| 0.015517 | 0.3519 | 0.1813 | 0.0000 | 0.7073 | 3 | 1 |
| 0.018644 | 0.5278 | 0.1660 | 0.2024 | 0.8531 | 4 | 1 |
| <0.020000 | | | | | 19 | 15 |
| 0.023529 | 0.7037 | 0.0879 | 0.5315 | 0.8759 | 20 | 1 |
| 0.024528 | 0.7407 | 0.0843 | 0.5754 | 0.9060 | 21 | 1 |
| **0.033962** | **0.7778** | 0.0800 | 0.6210 | 0.9346 | 22 | 1 |
| 0.049153 | 0.8148 | 0.0748 | 0.6683 | 0.9613 | 23 | 1 |
| 0.106061 | 0.8519 | 0.0684 | 0.7179 | 0.9858 | 24 | 1 |
| 0.174074 | 0.8889 | 0.0605 | 0.7703 | 1.0000 | 25 | 1 |
| 0.177049 | 0.9259 | 0.0504 | 0.8271 | 1.0000 | 26 | 1 |
| 0.268966 | 0.9630 | 0.0363 | 0.8917 | 1.0000 | 27 | 1 |

You can check this table to see that the results are the same as those of Helsel (2005).

# Example 4 – Validation of Group Comparison Statistics using Helsel (2005)

This section presents validation of the group comparison statistics. Helsel (2005) presents an example of results for comparing concentrations among three groups. These data are contained in the CONCENTRATION dataset.

The results for the overall test for determining difference in concentration patterns across groups is found on page 180. The log rank test results in a chi-square statistic of 16.2794 with probability level 0.000. The Gehan (Wilcoxon) test gives a chi-square statistic of 16.0761 with probability level 0.000. The results of the individual group comparison Gehan (Wilcoxon) tests are given on page 181. For comparing the low group to the medium group, the chi-square value is 0.68890 with probability level 0.407. For comparing the low group to the high group, the chi-square value is 7.09906 with probability level 0.008. For comparing the medium group to the high group, the chi-square value is 11.5275 with probability level 0.001.

These data can be run in this procedure to see that ***NCSS*** gets the same results. You may follow along here by making the appropriate entries or load the completed template **Example4** from the Template tab of the Nondetects Analysis window

**1 Open the CONCENTRATION dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **CONCENTRATION.S0**.
- Click **Open**.

**2 Open the Nondetects Analysis window.**

- On the menus, select **Analysis**, then **Nondetects**, then **Nondetects Analysis**. The Nondetects Analysis procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Nondetects Analysis window, select the **Variables tab**.
- Set the **Response Variable** to **Conc**.
- Set the **Nondetection (Censor) Variable** to **CNondet**.
- Set the **Group Variable** to **Group**.

**4 Specify the reports.**

- On the Nondetects Analysis window, select the **Reports tab**.
- Uncheck all reports except the **Logrank Test Summary** report.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

# Output

**Logrank Tests Section**
**Hypotheses**
H0: Distribution Functions are Equal Among Groups
HA: At Least One Group Distribution Functions Differs

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha = 0.05) |
|---|---|---|---|---|
| **Logrank** | **16.280** | **2** | **0.0003** | **Yes** |
| **Gehan-Wilcoxon** | **16.076** | **2** | **0.0003** | **Yes** |
| Tarone-Ware | 16.669 | 2 | 0.0002 | Yes |
| Peto-Peto | 16.359 | 2 | 0.0003 | Yes |
| Mod. Peto-Peto | 16.369 | 2 | 0.0003 | Yes |

**Multiple Pairwise Tests Section**
**Hypotheses**
H0: Distribution Functions are Equal
HA: Distribution Functions Differ

**Group Pair Tested: High vs. Low**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| **Logrank** | **7.360** | **1** | **0.0067** | **Yes** | **0.0200** | **Yes** |
| **Gehan-Wilcoxon** | **7.099** | **1** | **0.0077** | **Yes** | **0.0231** | **Yes** |
| Tarone-Ware | 7.282 | 1 | 0.0070 | Yes | 0.0209 | Yes |
| Peto-Peto | 7.385 | 1 | 0.0066 | Yes | 0.0197 | Yes |
| Mod. Peto-Peto | 7.378 | 1 | 0.0066 | Yes | 0.0198 | Yes |

**Group Pair Tested: High vs. Medium**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| Logrank | 11.398 | 1 | 0.0007 | Yes | 0.0022 | Yes |
| **Gehan-Wilcoxon** | **11.528** | **1** | **0.0007** | **Yes** | **0.0021** | **Yes** |
| Tarone-Ware | 11.931 | 1 | 0.0006 | Yes | 0.0017 | Yes |
| Peto-Peto | 11.454 | 1 | 0.0007 | Yes | 0.0021 | Yes |
| Mod. Peto-Peto | 11.470 | 1 | 0.0007 | Yes | 0.0021 | Yes |

**Group Pair Tested: Low vs. Medium**

| Test Name | Chi-Square | DF | Prob Level | Reject H0 (Alpha =0.05) | Bonferroni Adjusted Prob Level | Reject H0 (Alpha =0.05) |
|---|---|---|---|---|---|---|
| Logrank | 1.125 | 1 | 0.2888 | No | 0.8663 | No |
| **Gehan-Wilcoxon** | **0.689** | **1** | **0.4065** | **No** | **1.0000** | **No** |
| Tarone-Ware | 0.796 | 1 | 0.3723 | No | 1.0000 | No |
| Peto-Peto | 1.109 | 1 | 0.2923 | No | 0.8769 | No |
| Mod. Peto-Peto | 1.092 | 1 | 0.2961 | No | 0.8884 | No |

Notes:
The most commonly used test is the Logrank test.

You can check this table to see that the results are the same as those of Helsel (2005).

## Chapter 250

# Xbar R (Variables) Charts

## Introduction

This procedure generates various control charts useful for monitoring the average and variability of a process. The Xbar, EWMA, Moving Average, Individuals, Range, Standard deviation, and CUSUM charts are available. Various reports and a Capability Analysis are also available. A robust estimation method is available for automatically removing measurements that are outside the control limits from the calculation of the mean and standard deviation.

## Variables Control Charts

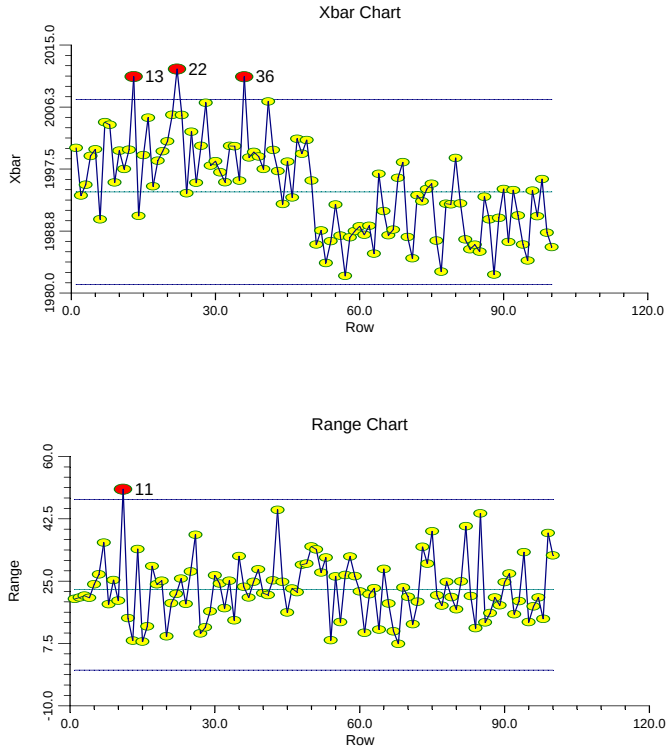
Suppose we have a scatter plot with a response variable on the vertical axis and a representation of time (such as hours, shifts, days, weeks, or months) on the horizontal axis. This scatter plot shows the nature of the response over time. For example, we might see trends, shifts, sudden jumps, and so on. If we add horizontal limit lines to the plot to indicate standards, the scatter plot becomes a *control chart*. When the plots fall inside these limits lines, the process yielding the response is said to be *in control*. When the process yields responses that are outside these limits, the process is said to be *out-of-control*.

The limit lines set a range of ‘normal behavior.’ They are based on past experience with the process and give a frame of reference for judging current outcomes. Because of natural variation in the process, the responses will not be exactly the same. They will bounce up and down. As long as the response stays within the limits, we need take no corrective action. However, once a measurement occurs outside the limits, we must investigate the cause and take appropriate corrective action.

Dr. Walter A. Shewart was the first to make the distinction between controlled and uncontrolled variation. While working at the Bell Telephone Laboratories in the 1920’s, he developed the control chart as a simple tool to separate the two forms of variation. Japan made extensive use of control charts. Their extraordinary success has led to the increasing use of control charts throughout the world. The power of control charts comes from their ability to signal the presence of assignable causes and provide a basis for improving the process.

Although there are many forms of control charts, they can be categorized as either *variables* or *attributes* control charts. Here, the term *variable* means that the data can take on any value. It does not have to be a whole number. A person's weight or height are examples of variables data. *Attributes*, on the other hand, are things that can be counted, such as the number of students in a class or the number of scratches on a new car. It doesn't make sense to talk about a half a scratch. The scratch either exists or it doesn't. If you can put things in categories such as good or bad, acceptable or not acceptable, then they are attributes data.

The most popular variables charts are usually referred to as *Xbar-R* charts. An *Xbar-R* chart is actually two plots, the *Xbar plot* and the *R plot*. The *Xbar* plot is a plot of averages on a control chart. The *R* plot is a plot of ranges of groups or responses across time. Often, these plots are shown together, with the range plot on the bottom. This allows both patterns to be studied together across time. An example of the *Xbar - R* chart is given in the following figure.




## Finding the Appropriate Control Limits

Once we understand that a control chart is simply a plot of some measurement across time with appropriate limits shown as horizontal lines, the only question is how to these limits should be determined. The answer to this question depends on the situation. For example, stock brokers use control charts routinely to determine when to buy and sell stocks. Unknowingly, they are using control charts. Each company sets its buy and sell limits in a different way, hoping to cash in on the movement of the stock.

In quality control work, these control limits are set to meet the needs of the people monitoring a process. By considering the past statistical behavior of the process, we can set the limits so that few false alarms are given. Typically, the statistical behavior of the process is represented by its

average and standard deviation. Statistical theory is used to set the limits so that, on the average, only about 3 in 1000 false alarms (saying a process is out-of-control when it is not) are generated.

The formulas given in a later section give the mathematical details on how to set the limits for different types of measurements.

## Comparison of Control Charts

Several types of control charts have been developed for the many situations that occur in practice. The first control charts were done by hand without the aid of computers or even calculators. Hence, techniques were developed that were easy to do by hand. With the advent of computers, more complex statistical techniques became available that have better properties.

How should we compare these techniques? What makes one charting procedure better than another? Various aspects of this question occur. For example, one way to compare two charting procedures is to investigate the average run length (ARL) after a known change in the process has occurred until an out-of-control signal is given. The ARL is the number of time periods that occur between the time a change actually occurs and the time an out-of-control signal is given by the chart.

It turns out that different charting procedures can have very different ARL's. The *Xbar* chart was developed to detect shifts in the process mean of about three sigmas (standard deviations). When a one sigma shift occurs, the ARL of the Xbar chart is about 6.3. The CUSUM chart, an alternative to the Xbar chart, has an ARL of only 3.2. When a mean shift of only one sigma is of interest to us, the CUSUM is obviously a better procedure.

There are many aspects to consider when choosing an appropriate charting procedure. We have already talked about the ARL. Another is the cost of sampling. Some procedures require larger subgroup sizes (the number of items measured at a particular point in time). Some procedures detect trends and patterns better than others. You will have to investigate which chart (or charts) is best for your situation.

Many books have been written about the pros and cons of the various control charts that are available. The following table gives a few of the advantages and disadvantages of the control charts that are available in this module.

## 250-4 Xbar R (Variables) Charts

| Chart | Focuses on | Subgroup Sample Size | Advantages | Disadvantages |
|---|---|---|---|---|
| Xbar | Average | Two and above | Does a good job at detecting sudden, large jumps in the process average.<br>Simple to understand.<br>Popular. Used often so there is a large body of knowledge about its use. | Slow to detect drifts in the average.<br>Not good at detecting small changes in the process average. |
| Individuals | Average | One | Does a reasonable job at detecting sudden jumps in the average.<br>Simple and popular. | Slow to detect drifts in the average.Not good at detecting small changes in the process average.<br>Relies heavily on the normality assumption. |
| EWMA | Average | One and above | Good at detecting slow shifts in the process average.<br>Can be effectively used with small group sizes. | Not responsive to sudden jumps.<br>Requires the setting of a subjective parameter. |
| Moving Average | Average | One and above | Good at detecting slow shifts in the process average.<br>Can be effectively used with small group sizes. | Not responsive to sudden jumps.<br>Not as effective as the EWMA chart. |
| CUSUM | Average | One and above | Detects small changes in the process average much sooner than the Xbar chart.<br>Is less likely to give false out-of-control signals. | Not as good at detecting larger jumps.<br>Somewhat more complicated. |
| R | Variability | Two and above | Good at detecting sudden jumps.<br>Easy to compute and understand. | Ignores a lot of information about the variability, especially when the subgroup size is large. |
| S | Variability | Two and above | Good at detecting sudden jumps.<br>Uses all information available about the variability contained in the data.<br>Theoretically optimal estimate of the variability in many situations. | Somewhat harder to compute and understand. |
| Moving Range | Variability | One | Only variability chart available when the subgroup size is one. | Ranges are no longer independent. Relies heavily on the normality assumption. |

# Formulas for Constructing Control Charts

Suppose we have *k* subgroups, each of size *n*. Let $x_{ij}$ represent the measurement in the $j^{th}$ sample of the $i^{th}$ subgroup. Often we set *n* to 5 and require *k* to be at least 25. Three statistics are routinely computed for each subgroup:

The subgroup mean

$$\bar{x}_i = \frac{\sum_{j=1}^{n} x_{ij}}{n}$$

the subgroup range

$$R_i = x_{(n)} - x_{(1)}$$

and the subgroup standard deviation

$$s_i = \sqrt{\frac{\sum_{j=1}^{n} \left(x_{ij} - \bar{x}_i\right)^2}{n-1}}$$

These three statistics are then plotted on the *Xbar chart*, the *R chart*, and the *s chart*, respectively.

## Estimating Sigma

Control limits must established for each of these statistics. These require an estimate of the process mean, $\mu_x$ (mu), and the process variability, $\sigma_x$ (sigma). Although a known estimate of $\mu_x$ may be supplied by the user, it is usually estimated by the average of the averages, 'x double bar' (also known as the grand mean):

$$\bar{\bar{x}}_i = \frac{\sum_{i=1}^{k} \bar{x}_{ij}}{k}$$

There are four methods available for estimating $\sigma_x$. First, it may be supplied by the user based on other information available to him. More frequently, however, it is estimated by one of the following methods:

### Method 1: Estimating Sigma from the Ranges

$$\hat{\sigma}_x = \frac{\bar{R}}{d_2}$$

where

$$\bar{R} = \frac{\sum_{i=1}^{k} R_i}{k}$$

$$d_2 = \frac{E(R)}{\sigma_x} = \frac{\mu_R}{\sigma_x}$$

Unfortunately, the calculation of *E(R)* requires the knowledge of the underlying distribution of the $x_{ij}$'s. Making the assumption that the $x_{ij}$'s follow the normal distribution with constant mean

and variance, we can derive values for $d_2$ through the use of numerical integration. These values are used in the program. It is important to note that the normality assumption is used and that the accuracy of this estimate requires that this assumption be valid.

## Method 2: Estimating Sigma from the Standard Deviations

$$\hat{\sigma}_x = \frac{\bar{s}}{c_4}$$

where

$$\bar{s} = \frac{\sum_{i=1}^{k} s_i}{k}$$

$$c_4 = \frac{E(s)}{\sigma_x} = \frac{\mu_s}{\sigma_x}$$

Again, the calculation of *E(s)* requires the knowledge of the underlying distribution of the $x_{ij}$'s. Making the assumption that the $x_{ij}$'s follow the normal distribution with constant mean and variance, we can derive values for $c_4$ from the following formula. It is important to note that the normality assumption is used and that the accuracy of this estimate requires that this assumption be valid. If the data come from the normal distribution, we can show that

$$c_4 = \sqrt{\frac{2}{n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

## Method 3: Estimating Sigma from the Mean Square Error

When the underlying data follow the normal distribution, the best estimate of $\sigma_x$ from a theoretical point of view is obtained by calculating the mean square error from a one-way ANOVA in which the subgroups are the treatments. The estimated value of $\sigma_x$ is

$$\hat{\sigma}_x = \sqrt{MSE}$$

Unfortunately, although this is the best estimate of $\sigma_x$, it is the least frequently used. Since charting techniques were done by hand, the complexity of the calculations discouraged most from using it. However, now that we have computer programs to do the calculations for us, there is no excuse for using one of the inferior estimates!

## Method 4: Estimating Sigma when n = 1

When $n$ is one, we cannot calculate $R_i$ or $s_i$ since these require at least two measurements. In this case, we could use the standard deviation of all $k$ measurements. Unfortunately, this method does not approximate the within-subgroup variation. Rather, it combines the within and the between subgroup variation. The common procedure is to use the ranges of successive pairs of observations. Hence, the range of the first and second is computed, the range of the second and third is computed, and so on. The average of these approximate ranges is used to estimate $\sigma_x$.

## Xbar Chart Limits

The lower and upper control limits for the Xbar chart are calculated using the formula

$$LCL = \bar{\bar{x}} - m\left(\frac{\hat{\sigma}_x}{\sqrt{n}}\right)$$

$$UCL = \bar{\bar{x}} + m\left(\frac{\hat{\sigma}_x}{\sqrt{n}}\right)$$

where $m$ is a multiplier (usually set to three) chosen to reduce the possibility of false alarms (signaling an out-of-control situation when the process is in control).

### Xbar Chart Limits when n=1

When the subgroup size is one, the control limits become

$$LCL = \bar{\bar{x}} - m\hat{\sigma}_x$$

$$UCL = \bar{\bar{x}} + m\hat{\sigma}_x$$

where $\hat{\sigma}_x$ is based on the moving ranges as described above.

## Range (R) Chart Limits

The lower and upper control limits for the range chart are calculated using the formula

$$LCL = R_e - md_3\hat{\sigma}_x$$

$$UCL = R_e + md_3\hat{\sigma}_x$$

where $m$ is a multiplier (usually set to three) chosen to reduce the possibility of false alarms and $d_3$ is a constant (which depends on $n$) which is calculated from the following relationship by numerical integration based on the assumption of normality.

$$d_3 = \frac{\sigma_R}{\sigma_x}$$

The value of $R_e$ is $\bar{R}$ if $\sigma_x$ is estimated from the data or by $d_2\sigma_x$ if $\sigma_x$ is supplied by the user.

### Range Chart Limits when n=1

The moving ranges of size two replace the usual ranges in the formulas above. All calculations remain the same after this substitution, except that there are only *k-1* ranges to plot.

## Sigma (S) Chart Limits

The lower and upper control limits for the sigma (standard deviation) chart are calculated using the formula

$$LCL = \hat{\sigma}_s - mf_1\hat{\sigma}_x$$

$$UCL = \hat{\sigma}_s + mf_1\hat{\sigma}_x$$

where $m$ is a multiplier (usually set to three), $f_1$ is a constant which, based on the assumption of normality, is given by the formula

$$f_1 = \sqrt{1 - c_4^2}$$

The value of $\hat{\sigma}_s$ is $\bar{s}$ if $\sigma_x$ is estimated from the data, or $c_4\sigma_x$ if $\sigma_x$ is supplied by the user.

## EWMA Chart Limits

The lower and upper control limits for the exponentially weighted moving-average (EWMA) chart are calculated using the formula

$$LCL_i = \bar{\bar{x}} - m\left(\frac{\hat{\sigma}_x}{\sqrt{n}}\right)\sqrt{\frac{\pi}{2-\pi}\left[1-(1-\pi)^{2i}\right]}$$

$$UCL_i = \bar{\bar{x}} + m\left(\frac{\hat{\sigma}_x}{\sqrt{n}}\right)\sqrt{\frac{\pi}{2-\pi}\left[1-(1-\pi)^{2i}\right]}$$

where $m$ is a multiplier (usually set to three) and $\pi$ is smoothing constant. The values plotted are obtained from the original $\bar{x}_i$'s using the exponential smoothing operation given by

$$e_i = \pi\bar{x}_i + (1-\pi)e_{i-1}$$

The value of $e_0$ is set to the grand mean.

Note that the values of limits change with each successive subgroup. Fortunately, the value of the radical stabilizes after $i$ passes five or six.

## Moving Average Chart Limits

The lower and upper control limits for the moving-average chart are calculated using the formula

$$LCL_i = \bar{\bar{x}} - m\left(\frac{\hat{\sigma}_x}{\sqrt{n_i w_i}}\right)$$

$$LCL_i = \bar{\bar{x}} + m\left(\frac{\hat{\sigma}_x}{\sqrt{n_i w_i}}\right)$$

where $m$ is a multiplier (usually set to three) and $w_i$ is the number of rows used in this average. Note that the value of $w_i$ changes during the first few rows and then stays constant. The values plotted are obtained from the original $\bar{x}_i$'s by taking the average of the last $w_i$ rows (including the current row).

## CUSUM Charts

The CUSUM chart has been shown to detect small shifts in the process average much quicker than the Xbar chart. In fact, it can be shown to be better than the Xbar chart in many ways. Until recently, however, a cumbersome procedure using the so-called V-mask was necessary. Now, however, a charting procedure similar to the Xbar chart is available.

In ***NCSS*** we use the CUSUM procedure presented by Ryan (1989). This procedure may be summarized as follows:

1. Calculate all statistics as if you were going to generate an Xbar chart.

2. Calculate the $z_i$ using the formula

$$z_i = \frac{\bar{x}_i - \bar{\bar{x}}}{\hat{\sigma}_{\bar{x}}}$$

3. Calculate the lower and upper cumulative sums as follows

$$S_{Li} = -\max\left[0, \left(-z_i - K\right) + S_{Li-1}\right]$$

$$S_{Hi} = \max\left[0, \left(z_i - K\right) + S_{Hi-1}\right]$$

4. Plot $S_{Hi}$ and $S_{Li}$ on a control chart. The control limits are chosen as plus or minus *h*. Often, *K* is set to 0.5 (for detecting one-sigma shifts in the mean) and *h* is set to 5.

5. When an out-of-control situation is detected, the corresponding sum is reset to an appropriate starting value. Usually, the starting value is zero. Occasionally, however, a "fast initial restart" (FIR) value of *h/2* is used.

## Runs Tests

The strength of control charts comes from their ability to detect sudden changes in a process that result from the presence of assignable causes. Unfortunately, the Xbar chart is poor at detecting drifts (gradual trends) in the process. For example, there might be a positive trend in the last ten subgroups, but until a value goes above the upper control limit, the chart gives no indication that a change has taken place in the process.

Runs tests are ways to check your control charts for unnatural patterns that are most likely caused by assignable causes. Years ago, statisticians referred to these patterns as runs and the term stuck, although today many people refer to them as "pattern tests" or "out-of-control" tests. The presence of any of these patterns means your process has probably changed, so you will need to find the problem and fix it.

The application of the runs tests is the same for all control charts. However, the interpretation of the results depends on which control chart you are using. We shall discuss some of the important differences as we go.

In order to perform the tests, the control chart is divided into six equal zones (three on each side of the centerline). Since the control limit is three sigma limits (three standard deviations of the mean) in width, each zone is one sigma wide and is labeled A, B, or C, with the C zone being the closest to the centerline. There is a lower zone A and an upper zone A. The same is true for B and C. The runs tests look at the pattern in which points fall in these zones.

We will now discuss each of the runs tests available in ***NCSS***.

### Test 1: Any Single Point Beyond Zone A

We have already discussed this runs test, although we did not call it that at the time. It is simply a point beyond the three-sigma control limit. This is the main test for an unnatural pattern. Since there is less than a 0.3% chance of this occurring naturally, it is a strong indication of an assignable cause.

In a range chart, a point above the upper-control limit indicates that the piece-to-piece variation has suddenly increased. Check for worn parts or variation in the raw material.

In the Xbar chart, a point beyond either control limit indicates a serious change in the process. Check points before and after the occurrence to see if this is an isolated case or part of a trend.

## Test 2: Two of Three Successive Points in Zone A or Beyond

This usually indicates a shift in the process average. Note that the two points have to be in the same Zone A, upper or lower. They cannot be on both sides of the centerline. The third point can be anywhere.

## Test 3: Four of Five Successive Points in Zone B or Beyond

This usually indicates a shift in the process average. Note that the odd point can be anywhere.

## Test 4: Eight Successive Points in Zone C or Beyond

All eight points must be on one side of the centerline. This is another indication of a shift in the process average.

## Test 5: Fifteen Successive Points Fall in Zone C on Either Side of the Centerline

Although this pattern might make you think that the variation in your process has suddenly decreased, this is usually not the case. It is usually an indication of stratification in the sample. This happens when the samples come from two distinct distributions having different means. Perhaps there are two machines that are set differently. Try to isolate the two processes and check each one separately.

## Test 6: Eight of Eight Successive Points Outside of Zone C

This usually indicates a mixture of processes. This can happen when two supposedly identical production lines feed a single production or assembly process. You must separate the processes to find and correct the assignable cause.

There are, of course, many other sets of runs tests that have been developed. You should watch your data for trends, zig-zags, and other nonrandom patterns. Any of these conditions could be an indication of an assignable cause and would warrant further investigation.

Two questions that inevitably arise in any discussion of runs tests are: “What is the probability that the runs tests will not detect a problem that really exists?” and “What is the probability that a runs test will tell me I have a problem when I really don’t?” The first question is difficult to answer because there are so many potential problems that it is virtually impossible to estimate the probability of occurrence for each of them. The best we can say is that over the years, control charts have been extremely successful in finding assignable causes. The companies that have used the charts and have found and corrected the problems that the charts have indicated have usually gained an edge on their competition.

The second question (that of false alarms) is easier to answer because the tests are structured to minimize the occurrence of false out-of-control conditions. With the exception of range charts with small sample sizes, the probability of getting a false alarm for any individual test is relatively small. It does exist, however, and for any given point on an Xbar chart with only common causes, there is about a 2% chance of getting a false indication of an out-of-control condition. The probability of a false alarm at any point on a range chart with a sample size of 5 is about 2.7%. These are acceptable probabilities for most people, but keep in mind that on an Xbar-R chart with a large number of samples (40 or more), you run a very good chance of having at least one false alarm.

If you are using a sample size of two, be careful! The probability of a false alarm at any point is nearly 23%. Any time you get an out-of-control signal, make sure you understand the cause. You can avoid a lot of false alarms by making the sample size four or five. You may also want to ignore the runs tests and concentrate on only those points that go beyond the three-sigma control limits.

## Capability Analysis

In all of our discussion of process performance, we have not yet mentioned the word *specification*. If you are manufacturing a product or even providing a service, you may be concerned about this omission. After all, it is the specification that the customer will check your product against, not the control limits. So you may be asking, "What good are control limits if they are not related to the specifications?" The answer to that question is the subject of this section.

After you have assured yourself that the process is stable and you have identified and removed all the assignable causes, the process will be in statistical control. Since the remaining variation is due to common causes only, the process is doing the best that can be expected. But is "the best" good enough? To find the answer to this question, you have to perform a *capability analysis*.

The basic idea of a capability analysis is to compare the process output with the specifications to determine whether the process can be expected to produce items that will be within the specification limits. If you enter the process specifications (or other requirements), **NCSS** will calculate the process capability for you.

In order to see how the process capability works, let's consider a typical hamburger restaurant. The owner is concerned about the weight of the hamburgers because if they are too small, the customers complain, and if they are too big, he loses money. He therefore directs the restaurant manager to make sure that all hamburgers are within one-half ounce of the advertised weight of 4 ounces. Hence the lower specification limit is 3.5 ounces and the upper specification limit is 4.5 ounces.

From previously created control charts, the manager knows that the process is in statistical control. If he made a histogram of the weights of a week's production of hamburgers, he would create a chart that would be close to the familiar bell curve. Suppose the mean hamburger weight is 4.1 ounces.

Before computers were available, it was often difficult to calculate the percentage of items that would fall outside of the specifications, so a number of shortcuts were developed. One of these shortcuts is called the *capability index*. Actually, there are two versions of the capability index, which are generally labeled *Cp* and *Cpk*. *Cp* evaluates the process spread relative to the specifications and *Cpk* evaluates the process location relative to the specifications.

Although the percentage of items produced outside of specification may be more meaningful for decision making, there are a lot of people that still prefer to use one of the capability indexes. If your main customer is one of these people, you will have to use an index yourself, so let's examine them briefly to see how they are used.

In order to find out what the process can do, we must first remove all assignable causes. If your process is not in statistical control, you cannot get a good estimate of the process capability. In order to get around this problem, you can make an estimate of the process capability without assignable causes by removing the data samples that fail any of the runs tests and then calculating the process capability. This is not as good an estimate as if the actual causes themselves were removed, but it is better than leaving the out-of-control points in.

*Cp* is the difference between the two specification limits divided by six-sigma. Mathematically, the equation is

$$Cp = \frac{USL - LSL}{6\sigma}$$

Large values of *Cp* are wanted, while small values are unwanted. The selection of the '6' in the denominator is so that when the grand mean is just equal to the target value and the underlying distribution is normal, a *Cp* = 1 indicates that only 0.27% of the process output will be outside specifications. This amounts to 27 out of 10,000.

Unfortunately, a *Cp* of one does not guarantee the 0.27%. All it guarantees is that when the normality assumption is correct, there will never be less than 0.27% outside specifications! As you can guess, the use of *Cp* has dropped off in more recent years. The problem is that it does not take into account the true center of the process output. If you look at the above formula, you realize that it is possible for *Cp* to be one, yet none of the process is inside specifications!

The *Cp* index has generally been replaced by the *Cpk* index because the *Cpk* index tries to take into account the process average as well as the process variability. The theoretical definition of *Cpk* is

$$Cpk = \frac{\min\left(USL - \mu_x,\ \mu_x - LSL\right)}{3\sigma}$$

or, upon using a little algebra,

$$Cpk = \left\{1 - \frac{\left|\mu_x - \frac{1}{2}\left(USL + LSL\right)\right|}{d}\right\} Cp$$

where $d = USL\text{-}LSL$.

When *Cpk* is greater than one, the process is said to be *capable*. An index smaller than one indicates a problem that needs attention. Because some processes produce an output that does not fit a Gaussian distribution, quality experts sometimes use a *Cpk* value of 1.33 to indicate a capable process. This more conservative value corresponds to plus or minus four-sigma limits instead of the more common three-sigma limits.

Actually, the most useful information from the capability study could be the estimate of the percentage of items that will fall outside the specification limits. As mentioned earlier, it used to be very difficult to calculate this number. But, with computers doing the work for us, it is now quite easy. **NCSS** calculates the percentage above the upper specification and the percentage below the lower specification and labels these values as "% Outside Spec." You can decide what is an acceptable threshold for this value in order for your process to be termed capable. If the process follows a Gaussian distribution, the *% Outside Spec* would have to be less than 0.3% total to correspond to *Cpk* being greater than one.

Because the calculation of *Cp* and *Cpk* depend on the characteristics of a normal distribution, it is important to know if your data are normal. Normal does not mean that your data are regular or standard. The question is really whether your individual measurements fit the bell shaped curve of a normal distribution. **NCSS** tests the distribution of underlying data to see how well it fits a normal (or Gaussian) distribution.

The test is called a chi-square test for normality. It makes a histogram by dividing the distribution into a number of regions and then counting the number of items in each region. It then compares this actual count with the number that is expected if the distribution is normal. The test uses these differences to calculate a chi-square value.

A small chi-square value (below some threshold value), means the data are well approximated by the normal distribution. A large chi-square value (above the threshold value) means that the data are not normally distributed. The chi-square threshold value depends on the number of data samples and the desired confidence level of the answer (the default is 95%). If the value of chi-square is above the threshold value, **NCSS** tells you to reject the hypothesis that the data follow a normal distribution.

If your data do not follow a normal distribution, be careful when interpreting the capability index. You could have more parts out of spec than what the index value would lead you to expect. Under these circumstances, you may want to use a value greater than one to indicate a capable process. Some people suggest using 1.33, which corresponds to limits of plus or minus 4-sigma.

Because some people prefer to use a confidence level other than 95%, **NCSS** provides both the actual chi-square value and the probability that a chi-square value equal to or greater than this could have come from normally distributed data.

The sensitivity of the chi-square test depends on the number of individual measurements in the data base. When there are fewer than 200 points in the test, the test may not properly reject the hypothesis of normality. However, the larger the data base, the more sensitive is the test, so that for large data bases (more than 200 points) even small deviations from normality will be detected.

Your control charts will be valid even if your data are not normal. This is because control charts are based on samples of measurements which are then averaged. As long as the sample size is three or greater, these averages will form a distribution that is close enough to Gaussian to make valid control charts regardless of the underlying individual distribution.

Although the chi-square test is of interest to statisticians, many times a quick look at a chart of the distribution will let you see if you have problems with the shape of the distribution. **NCSS** fits a normal curve to the histogram of your individual data to show you what your distribution looks like. From this, you can easily judge for yourself if there are any problems with your data.

## Issues in Using Control Charts

We would like to point out several decisions that must be made when using a control chart. We will not make these decisions for you. However, these are issues that you must deal with when adopting and using any control charting technique. The answers to the following questions are important. If you want help with these questions, we suggest that you obtain a book on the subject such as Ryan (1989) or Montgomery (1991). Such books will give you a much better background in the techniques of control charting.

### Subgroup Size

How many items per subgroup? Originally, four or five items were recommended. Nowadays, ten or twenty are not uncommon. What difference does it make? What about unequal subgroup sizes?

### Dealing with Out-of-Control Points

How do you deal with out-of-control points once they have been detected? Should they be included or excluded in the process average and standard deviation?

### Control Limit Multiple

I understand that most people use 3-sigma limits. What is so magic about 3? Are there situations where 3.1-sigma limits are more appropriate? How about 2-sigma limits?

### Startup Time

I understand that I should have about 25 periods of in-control readings before I pay much attention to the control chart. Is 25 subgroups enough for my situation? Should I have more or less?

### Normality Assumption

I hear a lot of discussion about the importance of having a measurement that is normally distributed. How important is this? How do I check this? How non-normal does a process have to be before I have to choose a different procedure?

### Runs Tests

I hear of all kinds of tests for pattern detection. Which runs tests (if any) are appropriate for my situation? Are they really useful, or do they just add extra work?

## Data Structure

The data in the table below illustrate how control chart data are entered into **NCSS**. Each row gives the five responses for a particular subgroup. It is the average and range of each row that is charted. These data represent fifty subgroups of five samples. The data are contained in the file named QATEST. Only the first eight rows of the data are shown here.

**QATEST dataset (subset)**

| S1 | S2 | S3 | S4 | S5 |
|---|---|---|---|---|
| 2 | | 3 | | 5 |
| 8 | 8 | 7 | 7 | 9 |
| 6 | 2 | 2 | 4 | 3 |
| 5 | 6 | 7 | 6 | 10 |
| 48 | 2 | 6 | 5 | 0 |
| 28 | 2 | 1 | 5 | 13 |
| 7 | 4 | 5 | 4 | 8 |
| | 0 | 5 | 7 | 3 |

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

## Variables Tab

This panel specifies the variables that will be used in the analysis.

### Variables

#### Data Variables

These are the variables to be analyzed, one for each sample. Each row represents a complete subgroup. For example, if your procedure is to take five samples per subgroup, you would enter the five values in five variables across a row.

If only one variable is given, **NCSS** automatically generates an individuals chart with a moving-range of size 2.

#### Label Variable

An optional variable containing row labels that you use to document your output. You can use dates (like Jan-23-95) as labels. Here is how. First, enter your dates using the standard date format (like 06/20/93). In the Variable Info screen, change the format of the date variable to something like *mmm-dd-yyyy* or *mm-dd-yy*. The labels will be displayed as labels. Without changing the variable format, the dates will be displayed as long integer values.

### Specify Rows in Calculations

#### Specification Method

This option specifies how the rows that are used in the calculations are specified.

- **All Rows**
  All rows are used.
- **First Row - Last Row**
  The first and last row is specified.
- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.
- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.
- **Row List**
  The rows used by in calculations are specified by the Row List box below.

#### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

### Row List

Specify sets of rows to be used in calculations. A separate set of calculations will be carried out for each set. Example (with three sets): 1-50, 75-150, 175-Last. Note that Specification Method must be set to Row List.

Rows that are not included in this list will still be plotted if they are included in the list of charted rows.

## Specify Rows in Charts

### Specification Method

This option specifies how the rows that are used in the charts are specified.

- **All Rows**
  All rows are used.
- **First Row - Last Row**
  The first and last row is specified.
- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.
- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

## Select Chart Attributes

### Mean Line

Specifies whether to display a horizontal line representing the mean on the charts.

### Primary Control Limits

Specifies whether to display horizontal lines representing the primary control limits on the charts.

### Secondary Control Limits

Specifies whether to display horizontal lines representing the secondary control limits on the charts.

### Trend Line

Specifies whether to display a trend line on the charts.

### Runs

Specifies whether to add a label identifying those subgroups which failed a particular runs test.

### Row Labels

Specifies whether to label each row along the horizontal axis using the values in the Label Variable or the row number.

### Spec Limits on Chart

Specifies whether to display horizontal lines representing the specification limits on the charts.

### Spec Limits on Histogram

Specifies whether to display lines representing the specification limits on the histogram.

### Zones

Specifies whether to display horizontal lines representing the six horizontal zones.

### Individual Data

Specifies whether to display the individual data values. Typically, you would not show the individual values on a control chart since the control limits are for the average values, not the individual values. Occasionally, you might want to display the individual values with the specification limits on a time plot. This option will let you do that. Again, the control limits do not apply to the data values.

### Use Runs Tests

Specifies whether to use the runs tests in determining out-of-control points.

### Label Out-of-Control Rows

Specifies whether to label rows that fall outside the control limits. If a Label Variable is used, the label specified there is used. Otherwise, the row number of the out-of-control point is given.

## Options Tab

The next few options determine the type of chart that you want displayed.

### General Chart Options

#### Primary Multiplier

This option specifies the multiplier of sigma for the primary control limits. Usually, the famous 3-sigma limits are desired, so the multiplier is 3.

### Secondary Multiplier

This option specifies the multiplier of sigma for the optional, secondary control limits. Usually, the secondary limits are ignored by setting this value to 0. Occasionally, a value of 2 is used.

### Label Mean and Control Limits

Specifies whether to show the values of the mean (center) and controls limits on the right of the chart. You can also add optional headings like LCL= or Mean= with the option "Yes-Labels."

### Robust Options

You can have ***NCSS*** scan your data and remove out-of-control subgroups from the calculation of the mean and standard deviation. Usually, you would perform this manually by repeatedly removing out-of-control subgroups. Occasionally, you will want to obtain the final result without the manual intervention. These options define the automatic procedure that you want to use.

Caution: Accepted SPC procedure is to only remove out-of-control subgroups from the calculations if an assignable cause for the out-of-control situation can be found and corrected. If the problem cannot be found and corrected, you should not remove the out-of-control subgroup. Hence, you should be careful about using this automated procedure, since it does not require you to find assignable causes for the out-of-control points.

### Robust Iterations

This option specifies the number of robust iterations (cycles through the data looking for out-of-control subgroups) that you want. Usually, one or two should be sufficient.

If you want to skip the robust estimation entirely, enter a 0 here.

### Robust Multiplier

This option specifies a control limit multiplier to be used during the robust estimation. Usually, you would enter a 3 here. Occasionally, you might want to adjust this value slightly.

### Rows Skipped

If you print out individual row labels along the horizontal axis and you have many rows, the labels may over-write each other. This option lets you skip *x* number of labels. For example, if you only wanted to display every other label, you would enter a 1 here. If you only wanted to display every fifth label, you would enter a 4 here.

## Xbar Chart Options

### Fixed Xbar Type

A fixed value for Xbar (the mean) can be input as a constant or as the first row in a specified variable. This option specifies where to find the fixed mean value. It specifies which of the following two places to look for the fixed mean value.

### Xbar Constant

This option lets you specify a fixed value for Xbar (the mean). It requires the appropriate selection of Constant in the last option.

### Xbar Variable

This option lets you specify a fixed value for Xbar (the mean). To use this, you would enter the fixed value of the mean as the first value in this variable on the database. This might be convenient when several databases (each with different means) are being run. It requires the appropriate selection of Variable in the above option.

## R (Range) Chart Options

### Range Chart Type

This option specifies whether an R chart (based on ranges) or an S chart (based on standard deviations) should be used. The R chart is more popular because the ranges are easier to compute by hand, but the S chart has better theoretical properties.

### Sigma From

This option specifies the method used to estimate sigma.

- **Data**

  Sigma is estimated from either the average of the subgroup ranges or the average of the subgroup standard deviations, depending on the selection in the Range Chart Type option.

- **Mean Square Error**

  Sigma is estimated using the mean square error. If the data follow the normal distribution, this estimate has the best theoretical properties.

- **Fixed Value**

  Sigma is not estimated from the data. Rather, the value of sigma is specified by the user. Use this if you want to use a certain value of sigma.

### Fixed Sigma Type

A fixed value for sigma (the standard deviation) can be input as a constant or as the first row in a specified variable. This option specifies where to find the fixed sigma value. It specifies which of the following two places to look for the fixed sigma value.

### Sigma Constant

This option lets you specify a fixed value for sigma (the standard deviation). It requires the appropriate selection of Constant in the last option.

### Sigma Variable

This option lets you specify a fixed value for sigma (the standard deviation). To use this, you would enter the fixed value of the sigma as the first value in this variable on the database. This might be convenient when several databases (each with different sigmas) are being run. It requires the appropriate selection of Variable in the above option.

## EWMA Option

### EWMA Parameter

This specifies the value of the smoothing parameter, $\pi$, in the EWMA chart. Typically, a value between 0.15 and 0.30 is used.

## Moving Average Option

### Moving Average Width

This specifies the number of rows averaged in each moving average. The moving average used the designated number of rows, going back from the current row. The xbar value of the current row is included in the moving average. At the beginning of the series, a reduced number of rows is used (since they are all that is available).

## Spec Limits

### Lower Specification Limit

This option lets you specify the optional lower specification limit for display on your charts and for use in the capability analysis.

### Upper Specification Limit

This option lets you specify the optional upper specification limit for display on your charts and for use in the capability analysis.

### Target Specification

This option lets you specify the optional target specification for display on your charts.

## Capability Section Option

### Alpha Level

This option specifies the value of alpha used in the confidence limits that are displayed in the capability analysis. Typically, this value is set to 0.05.

# Reports Tab

The following options control the format of the reports.

## Specify Reports

### Chart Summary Section - Capability Analysis Section

Each of these options controls the display of the corresponding report.

## Specify Charts

### Xbar Chart - Capability Histogram

Each of these options controls the display of the corresponding chart.

## Report Options

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Variable Names

This option lets you select whether to display variable names, variable labels, or both.

### Page Title

This option specifies a title to appear at the top of each page.

### Plot Subtitle

This option specifies a subtitle to appear at the top of each plot.

## Xbar Charts Tab

This panel sets the options used to define the appearance of the xbar chart.

### Vertical and Horizontal Axis

#### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

#### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

### Xbar Chart Settings

#### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

# R Charts Tab

This panel sets the options used to define the appearance of the range or s chart.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## R Chart Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Show Lower Limit

This check box indicates whether to display the lower limit on the range chart. Often, only the upper limit is of interest.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

# CUSUM Tab

This panel sets the options used to define the appearance of the CUSUM chart.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## CUSUM Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Threshold Limit (h)

Specify value of the threshold limit, *h*. Typically, a 5 is used here.

### Reference Value (K)

Specify value of the reference value, *K*. Typically, a 0.5 is used here.

### Restart Method

Specify the method used to restart the sum once an out-of-control signal has been received. Usually, the sum is reset to zero. Occasionally, you might want to restart the sum at *h/2*, by selecting the *FIR* option.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Label (Y and X)

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Minimum (Y)

This option specifies the minimum value displayed on the vertical (Y) axis. If left blank, it is calculated from the data.

### Maximum (Y)

This option specifies the maximum value displayed on the vertical (Y) axis. If left blank, it is calculated from the data.

### Tick Marks - Ref. Numbers (Y and X)

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Y Major Ticks - Y Minor Ticks

These options set the number of major and minor tickmarks displayed on the vertical axis.

### Y Grid Lines

This check box indicates whether the grid lines that emanate from the vertical axis should be displayed.

## Histogram Tab

This panel sets the options used to define the appearance of the histogram.

### Vertical and Horizontal Axis

#### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

#### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## Histogram Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

Note that this plot is a special version of the scatter plot, so it uses a scatter plot style file, not the histogram style file as you might think.

### Show Normal Line

Specify whether to display the normal (gaussian) density line on the histogram.

### Number Line Values

Specify the number of increments to use to make the normal density line.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

# Symbols - Lines Tab

This panel specifies the plotting symbols and lines used on the charts.

## Symbols

### Out of Control

Specify the symbol used to display the out-of-control points. Click the button on the right to display the symbol modification window.

### In Control

Specify the symbol used to display the in-control points. Click the button on the right to display the symbol modification window.

### Data Value

Specify the symbol used to display the data values. Click the button on the right to display the symbol modification window.

## Lines

### Connecting Line - Histogram Line

These options specify the color, width, and style of the various lines that make up the control chart.

# Storage Tab

The options on this panel control the automatic storage of the means and ranges on the current database.

## Storage Variables

### Store Means in Variable

You can automatically store the means of each row into the variable specified here.

Warning: Any data already in this variable is replaced. Be careful not to specify variables that contain important data.

### Store Ranges (Sigmas) in Variable

You can automatically store the range (or standard deviation) of each row into the variable specified here. The choice of whether the range or the standard deviation is stored in this variable depends on which type of R chart was selected.

Warning: Any data already in this variable is replaced. Be careful not to specify variables that contain important data.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running an Analysis using Xbar R Charts

This section presents an example of how to run an analysis. The data used are found in the QATEST database. We will analyze the variables S1 through S5 on this database. In order to do a capability analysis, we will set the specification limits to 1.0 and 14.0. (Note that these limits are not necessary for the *Xbar-R charts*.)

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Xbar R (Variables) Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Xbar R (Variables) Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Xbar R (Variables) Charts**. The Xbar R (Variables) Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Xbar R (Variables) Charts window, select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** through **S5** from the list of variables and then click **Ok**. “S1-S5” will appear in the Data Variables box.

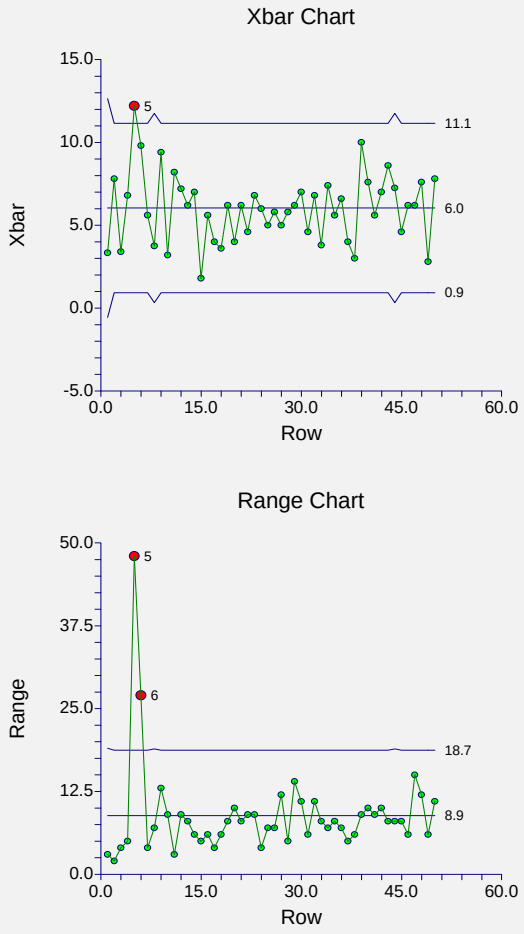
**4 Specify the specification limits.**

- On the Xbar R (Variables) Charts window, select the **Options tab**.
- Enter **1.0** in the **Lower Spec Limit** text box.
- Enter **14.0** in the **Upper Spec Limit** text box.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Xbar-R and Range Charts




This plot displays an *Xbar chart* on the top and an *R (range) chart* on the bottom. The overall mean (center-line) and 3-sigma limits are shown. These limits widen a little at the end because a missing value in batch 44 caused the sample size to be reduced from five to four. ***NCSS*** automatically adjusts for this change in sample size.

Notice that row 5 is outside the 3-sigma limits on the *Xbar chart*. The next report gives the numerical details of the charts and lists those rows that failed at least one of the control tests.

## Control Limits Section

**Control Limits Section**

| Control Limit | Xbar | Range | Sigma |
|---|---|---|---|
| Lower | 0.9261222 | -1.287481 | 0 |
| Upper | 11.14705 | 19.00748 | 0 |

This section displays the values of the lower and upper control limits for the two charts. Note that since the *S chart* was not run in this example, the control limits are both zero.

# Estimation Summary Section

**Estimation Summary Section**

| Estimate of Sigma | Mean | Range | Sigma-bar | Sigma |
|---|---|---|---|---|
| User Specified | 0 | 0 | 0 | 0 |
| Mean Square Error | 6.036585 | 8.86 | 3.611263 | 4.618599 |
| Ranges* | 6.036585 | 8.86 | 3.611263 | 3.809114 |
| Standard Deviations | 6.036585 | 8.86 | 3.611263 | 3.841828 |

| | |
|---|---|
| Number of Rows | 50 |

This report gives the numerical details of the Xbar-R chart analysis. We'll now define each of the numbers appearing on the report.

### Estimate of Sigma

There are four different estimates of sigma ($\sigma_x$) that may be selected. This report gives each of them along with the values that were used in their calculation. Note that the actual estimate of sigma is given in the last column under the heading "Sigma."

**User Specified** This is used if a value of sigma was specified by the user.

**Mean Square Error** This is the estimate of sigma that is calculated from the mean squared error of the data. This is the most efficient estimate of sigma when a process is in control in that it makes the best use of all information.

**Ranges** This estimate of sigma is based on the average of all the ranges. This is the most popular estimate of sigma. The star (*) by the word "Ranges" indicates that this is the estimate that was used in the current chart.

**Standard Deviations** This estimate of sigma is based on the average of all sample standard deviations.

### Mean

This gives the grand mean: the average of the subgroup xbars.

### Range

This is the average of the ranges.

### Sigma-bar

This is the average of the standard deviations.

### Sigma

This is the estimated value of $\sigma_x$.

# Out-of-Control List

**Out-of-Control List**

| Row | Mean | Range | Row Label | Reason |
|---|---|---|---|---|
| 5 | 12.2 | 48 | 5 | Xbar: beyond control limits |
| 6 | 9.8 | 27 | 6 | Range: beyond control limits |

This report provides a list of the rows that failed one of the runs tests (including being outside the control limits. The Reason column names the particular runs test that was failed.

# Capability Analysis Section

**Capability Analysis Section**

| Parameter | Lower | Center | Upper |
|---|---|---|---|
| 3-Sigma Limits | -5.390758 | 6.036585 | 17.46393 |
| 4-Sigma Limits | -9.199872 | 6.036585 | 21.27304 |
| Specification Limits | 1 | | 14 |
| Specification z-Values | -1.322246 | | 2.090621 |
| Percent Outside Specification | 3.252033 | | 2.439024 |
| Capacities | 0.440749 | | 0.696874 |
| Cp Index | 0.518452 | 0.568811 | 0.619112 |
| Cpk Index | 0.383108 | 0.440749 | 0.498389 |

Count = 246 Sigma = 3.809114 Alpha Level = 0.050000

This report provides a capability analysis of the data. The aim of a capability analysis is to test whether the process is capable of meeting the design specifications.

### 3-Sigma Limits

These are the estimated values of the 3-sigma limits. The grand mean, x double bar, is given in the center column.

### 4-Sigma Limits

These are the estimated values of the 4-sigma limits.

### Specification Limits

These are the specification limits that you entered.

### Specification z-Values

These are the z-values of the specification limits calculated using the formula:

$$z_{spec} = \frac{spec - \bar{\bar{x}}}{\hat{\sigma}_x}$$

Here ‘spec’ refers to the upper or lower specification limit.

### Percent Outside Specification

This is the percent of the individual sample values that were outside the specification limits. The first number is the percent that are less than the lower specification limit. The second number is the percent that are above the upper specification limit.

### Capacities

These are the absolute values of the above z-values divided by three. These values are used in the calculation of Cpk.

### Cp

This is the difference between the two z-values divided by six. When this number is greater than one, the process is said to be ‘capable.’

$$Cp = \frac{USL - LSL}{6\hat{\sigma}_x}$$

Also included are upper and lower confidence limits for the *Cp* value using the following equations

$$Cp_{lower} = Cp\sqrt{\frac{\chi^2_{n-1,\,\alpha/2}}{n-1}}$$

$$Cp_{upper} = Cp\sqrt{\frac{\chi^2_{n-1,1-\alpha/2}}{n-1}}$$

### Cpk

This is the minimum of the two *Capacities*. When this measure is greater than one (some people use 1.33) the process is said to be 'capable.'

$$Cpk = \frac{\min\left(USL - \bar{\bar{x}}, \bar{\bar{x}} - LSL\right)}{3\hat{\sigma}_x}$$

Also included are upper and lower confidence limits for the *Cpk* value using the following equations

$$Cpk_{lower} = Cpk - z_{1-\alpha/2}\sqrt{\frac{n-1}{9n(n-3)} + \left(\frac{Cpk^2}{2n-6}\right)\left(1+\frac{6}{n-1}\right)}$$

$$Cpk_{upper} = Cpk + z_{1-\alpha/2}\sqrt{\frac{n-1}{9n(n-3)} + \left(\frac{Cpk^2}{2n-6}\right)\left(1+\frac{6}{n-1}\right)}$$

### Count

This is the number of values used in the analysis. Normally, all values are included in this analysis. If you want to restrict the values used to those subgroups that are in control, you must use the runs tests and robust estimation procedures.

### Sigma

This is the estimated standard deviation of the underlying process.

### Alpha

This is value of α used in the confidence intervals of *Cp* and *Cpk*.

## Frequency Distribution and Normality Test

**Frequency Distribution and Normality Tests**

| Lower Boundary | Upper Boundary | Actual Count | Normal Count | Diff. Count | Actual Percent | Normal Percent | Diff. Percent | Chi-Sqr Amount |
|---|---|---|---|---|---|---|---|---|
| | -3.486201 | 0.0 | 1.5 | -1.5 | 0.0 | 0.6 | -.6 | .00 |
| -3.486201 | 0.322913 | 8.0 | 14.9 | -6.9 | 3.3 | 6.1 | -2.8 | 4.33 |
| 0.322913 | 4.132028 | 87.0 | 59.5 | 27.5 | 35.4 | 24.2 | 11.2 | 12.75 |
| 4.132028 | 7.941143 | 72.0 | 94.2 | -22.2 | 29.3 | 38.3 | -9.0 | 5.23 |
| 7.941143 | 11.75026 | 63.0 | 59.5 | 3.5 | 25.6 | 24.2 | 1.4 | .21 |
| 11.75026 | 15.55937 | 13.0 | 14.9 | -1.9 | 5.3 | 6.1 | -.8 | .01 |
| 15.55937 | | 3.0 | 1.5 | 1.5 | 1.2 | 0.6 | 0.6 | .00 |
| Total | | 246.0 | 246.0 | 0.0 | 100.0 | 100.0 | 0.0 | 22.53 |

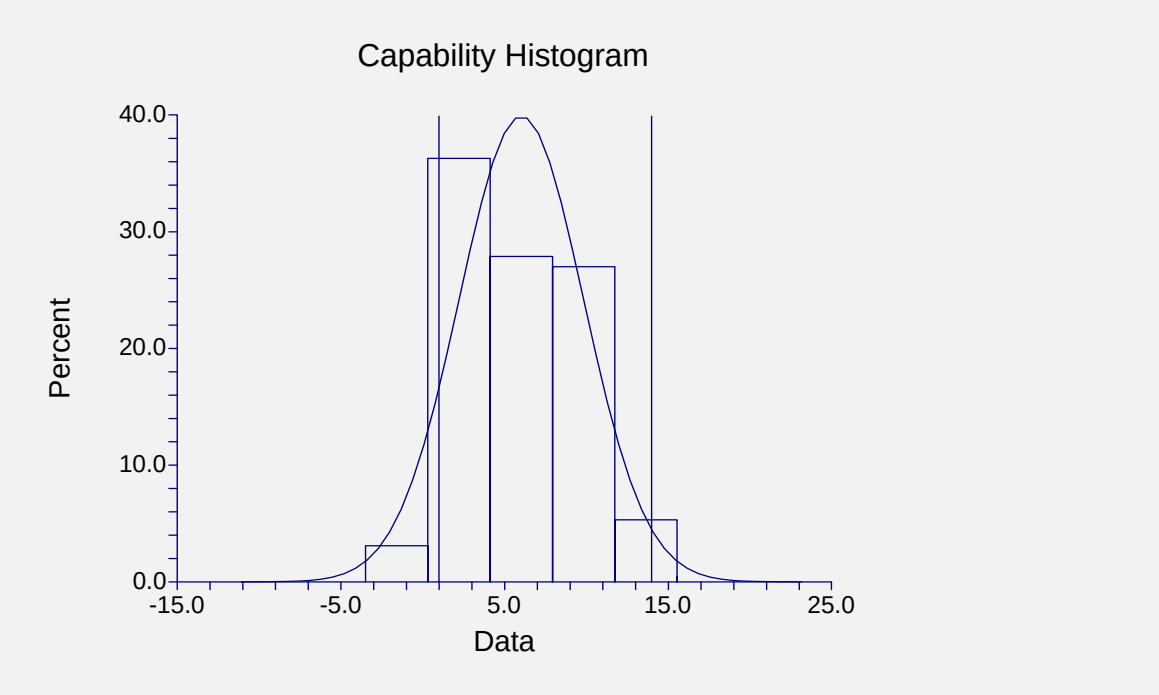
Normality: Chi-Square = 22.53 Prob Level = 0.000013. Normality hypothesis is rejected.
Normality: Shapiro-Wilk = 0.77 Prob Level = 0.000000. Normality hypothesis is rejected.
Normality: Anderson-Darling = 5.79 Prob Level = 0.000000. Normality hypothesis is rejected.

This table summarizes the multinomial Chi-square test for normality as applied to the individual data points. It tests the normality of the underlying data, not of the *Xbar's*. The Chi-square amounts are displayed so that you can see where the largest contributions to the Chi-square values come from. Note that the Chi-square test requires expected group sizes of 5 or more. Because of this, **NCSS** automatically combines groups that have expected counts less than 5.

The Chi-square test and rejection probability is shown at the bottom of the report.

In addition to the Chi-square normality test, the Shapiro-Wilk and the Anderson-Darling normality tests are also displayed (see discussion in the Descriptive Statistics chapter for details). These tests are recommended over the older Chi-square test.

## Capability Histogram




This chart displays a histogram of the data with a normal curve overlaid. It allows you to visually check whether the data follow the normal distribution. Note that even if the data are not normally distributed, if the subgroup size is at least five, the *Xbar chart* is valid. This is based on the central limit theorem, which states that averages of samples are approximately normally distributed regardless of the underlying individual distribution.

# Example 2 – Individuals and Moving Range Charts

We will now run an example of an *Individuals* and *Moving Range Chart.* These are run on a single variable, so we will run this example on S1.

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Xbar R (Variables) Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Xbar R (Variables) Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Xbar R (Variables) Charts.** The Xbar R (Variables) Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Xbar R (Variables) Charts window, select the **Variables** tab.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** from the list of variables and then click **Ok**. "S1" will appear in the Data Variables box.

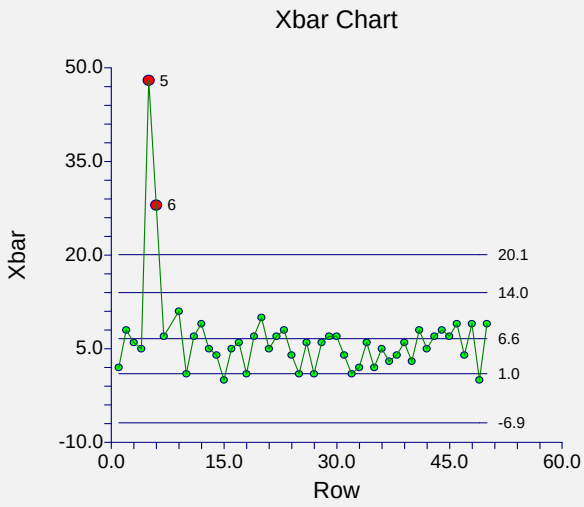
**4 Specify the specification limits.**

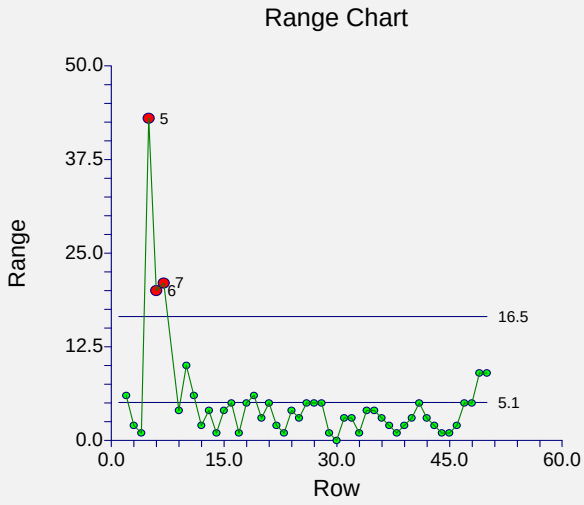
- On the Xbar R (Variables) Charts window, select the **Options** tab.
- Enter **1.0** in the **Lower Spec Limit** text box.
- Enter **14.0** in the **Upper Spec Limit** text box.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Output




The charts and reports of the Individuals and Moving Range charts appear about the same as before, so we will not repeat all the definitions here.

We wish to emphasize again that since the individuals chart is not based on averages, it cannot use the central limit theorem to assume normality. Instead, you must check the normality of your data very carefully before you can validate the use of this method.

# Example 3 – EWMA Charts

We will now run an example of a EWMA chart using the variables S1 through S5.

You may follow along here by making the appropriate entries or load the completed template **Example3** from the Template tab of the Xbar R (Variables) Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Xbar R (Variables) Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Xbar R (Variables) Charts**. The Xbar R (Variables) Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Xbar R (Variables) Charts window, select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** through **S5** from the list of variables and then click **Ok**. "S1-S5" will appear in the Data Variables box.

**4 Specify which reports and charts.**

- On the Xbar R (Variables) Charts window, select the **Reports tab**.
- Check **EWMA Xbar Chart**.
- Check **EWMA R (Range) Chart**.
- All of the other reports should not be checked.

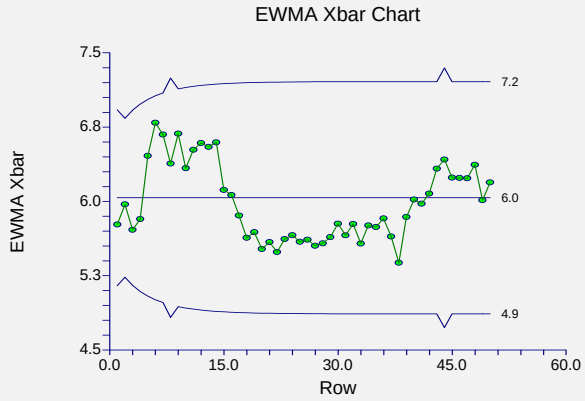
**5 Show both limits.**

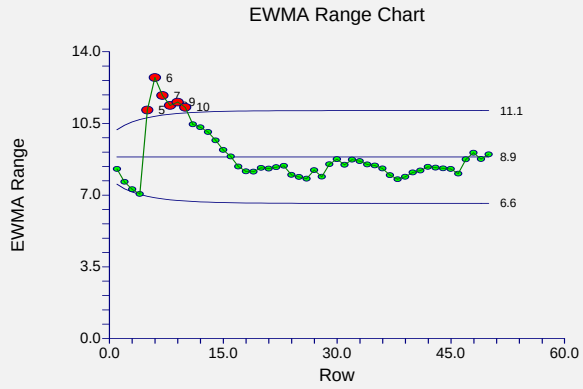
- On the Xbar R (Variables) Charts window, select the **R Charts tab**.
- Check the **Show Lower Limit** option.

**6 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Output




The charts and reports of the EWMA procedure appear about the same as before, so we will not repeat all the definitions here. Note, though, the characteristic widening of the first few limits.

# Example 4 – CUSUM Charts

We will now run an example of a CUSUM chart using the variables S1 through S5.

You may follow along here by making the appropriate entries or load the completed template **Example4** from the Template tab of the Xbar R (Variables) Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Xbar R (Variables) Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts,** then **Xbar R (Variables) Charts**. The Xbar R (Variables) Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Xbar R (Variables) Charts window, select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** through **S5** from the list of variables and then click **Ok**. "S1-S5" will appear in the Data Variables box.

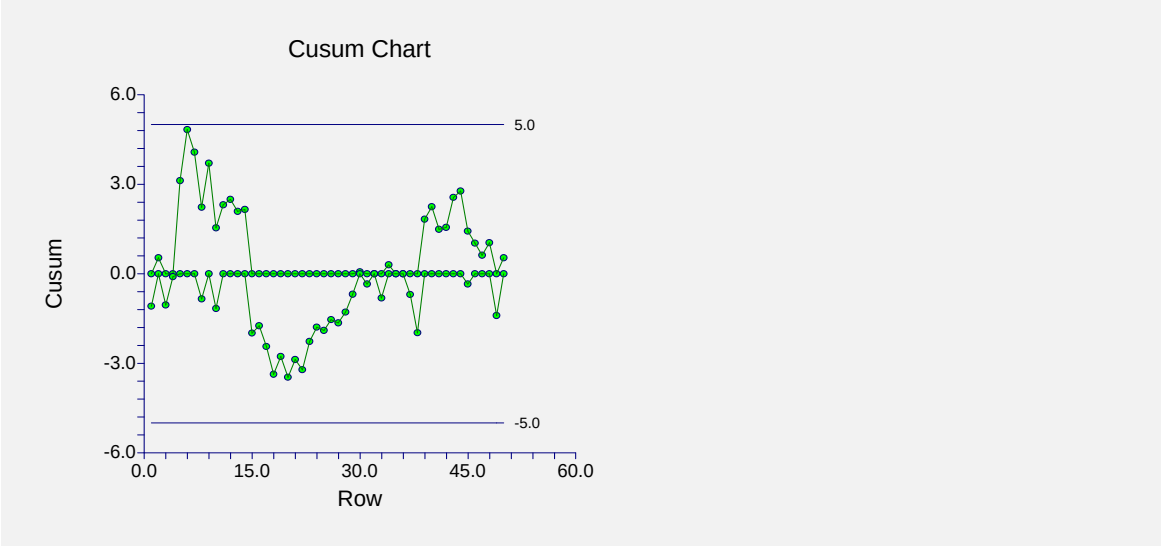
**4 Specify which reports and charts.**

- On the Xbar R (Variables) Charts window, select the **Reports tab**.
- Check **CUSUM Chart**.
- All of the other reports should not be checked.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Output




The charts and reports of the CUSUM procedure appear about the same as before, so we will not repeat all the definitions here.

# SPC Fundamentals

This section gives a brief introduction to SPC. We will explain what SPC is, how to use it, and what you can gain from it. If you need more detailed information on SPC, you should consult one of the many textbooks on SPC. We recommend Ryan (1989), Montgomery (1991), or DeVor (1992).

The use of statistics to help understand and control processes is certainly not a new idea. SPC has long been recognized as an extremely powerful tool to diagnose problems in quality and productivity. Recently, however, two events have occurred to enhance SPC's popularity. First, consumers throughout the world are becoming less tolerant of poor quality. A company that gains a reputation for good quality can actually charge a premium for its products and still gain market share. It also has a much easier time introducing new products that sell. Second, recent advances in computer hardware and software have simplified or eliminated the tedious calculations that formerly had to be done by hand.

SPC is based on the premise that there is variation in everything we do. There is no perfect process. Even the most sophisticated, numerically-controlled, machine varies slightly each time it repeats a process. The power of SPC comes from its ability to determine how much of the variation is from natural (common) causes that are inherent to the process and how much is from external (assignable) causes. This vital information helps determine the adequacy of the process and provides insight into how the process can be improved.

## Variation

Before we plunge into the depths of SPC, we need to discuss the concept of variation. Remember, we said earlier that there is variation in everything. No matter how hard we try to make things identical, they always turn out a little different. For example, a restaurant may advertise a "quarter-pound" hamburger, but any specific hamburger will probably not weigh exactly 4 ounces.

If we were to select 100 hamburgers at random and weigh them, we would find that each has a slightly different weight. If we were to construct a histogram of the weights of the hamburgers, we would find that most of them are very close to four ounces, but some are a little larger and some a little smaller. This pattern of weights is called a *frequency distribution*. By studying the shape of the distribution, we can gain a lot of useful information about the process.

One particular pattern that seems to appear quite often in nature is called the normal or Gaussian distribution. It is called the Gaussian distribution after Karl Gauss, a German mathematician who wrote an equation to describe the pattern. Because of its distinctive bell shape, it is often referred to as a "bell curve." The term *normal* is misleading because it implies that it is the most common or acceptable pattern.

Although patterns similar to the Gaussian distribution are very common, they are by no means the only patterns that you will encounter. The process variation can form some pretty strange shapes. Frequency distributions are often characterized by their location and spread.

The *location* of the distribution is often described by an average value called the *mean*. If the distribution is normal, the mean is the location of the peak value of the curve. If we were to randomly select one item from a normal distribution, it would be just as likely to fall above the mean as below it.

The *spread* or *variation* of a distribution is measured by what mathematicians call a standard deviation. If the distribution is normal, one standard deviation is the distance from the mean along the X-axis that includes about 34% of the data values in the distribution. If we measure one standard deviation on each side of the mean, we would include about 68% of the total data values. Two standard deviations on each side would take in over 95% and three standard deviations would include essentially all (more than 99%) of the data values in the distribution.

Another measure of the distribution spread or variation is the *range*. This is simply the separation between the largest and smallest value in the distribution. In other words, find the largest value and subtract the smallest value from it to get the range. Although the range works well when we're dealing with only a few data values, it does not describe the spread of a large number of items very well. For a normal distribution, the theoretical range is infinite.

### Causes of Variation

The causes of data variation can be grouped into two categories, *common* and *assignable*. *Common causes* are those that we cannot control unless we change the process itself. They are random in nature and an inherent part of the process. They are sometimes referred to as natural or system causes.

*Assignable causes*, on the other hand, are those that can be linked to a specific, correctable phenomenon. Variations due to assignable causes (sometimes called special causes) tend to distort the usual distribution curve and prevent the process from operating at its best. If assignable causes are present, the shape of the process distribution will vary with time.

## Steps to Create a Variables Control Chart

A control chart is only one of many tools of an effective quality program. In order to get the most out of a control chart, other elements of quality improvement—such as strong management, commitment to quality, and willingness to change—must also be in place.

In order to create effective Xbar-R charts, you should follow these steps:

## 1. Decide What to Measure.

This may sound easy, but watch out for traps. Remember, your goal is to control and, perhaps, improve the process. Therefore, you should measure something that is significant to the process. When it changes, it should signal a change in the process that is important to know. For example, a slight change in paint thickness may not be as important as a change in the yield strength of a metal. You can't afford to measure everything, so you should concentrate your measurement efforts where they will do the most good. You can find some good candidates for measurement by considering known problem areas. Look at scrap or rework, for example, and measure the parameters that seem to have the most effect on the problem.

You should also choose something that is relatively easy to measure. The appearance of your product is important but may be difficult to measure directly. Instead, you may have to decide what characteristics add to the appearance and then measure those. For example, the surface roughness and color both effect appearance and can be measured.

## 2. Gather Data.

First, you need to decide how many items (the sample size) you will measure in each sample and how often you will measure them. The larger the sample size, the more sensitive the control chart will be to a shift in the mean. Usually, a sample size of three to five is adequate. If you measure fewer than this, you may not be able to detect significant shifts in the process mean. Samples larger than five or six are more expensive, so you must balance cost with sensitivity. It is extremely important to choose the sample size so that all the pieces in the sample are produced under the same production conditions and within a very short time interval so that any variation within the sample is due only to common causes.

There is another reason why the sample size is important. Control charts are based on the properties of Gaussian distributions. Unfortunately, the distribution of items you are working with may not be remotely Gaussian. But when you sample several items from a distribution of any shape and plot the average of each sample, the distribution of the averages resembles a Gaussian distribution. The larger the sample size, the closer this distribution of means approaches to the Gaussian. Shewart found that even for sample sizes as small as four, the distribution of sample means was close to Gaussian.

How often you sample (the sample frequency) depends on the process you are trying to control. You need to sample often enough to catch changes as they occur, but sampling too often can be expensive. If you have a good idea of what can change and how quickly it can change, you can use this information to determine how often to take your samples. Unfortunately, many times we don't know what to expect, so initially you may want to sample at short time intervals to catch any quick changes. If you find the process is relatively stable, you can lengthen the sampling interval. For well behaved processes, sampling frequencies of one per hour, two per shift, or even one per day may be adequate.

People often ask how many samples are needed to detect a problem. There are two aspects to this question. The first is related to the process. You need to continue sampling and testing for a sufficient period to allow any assignable causes that might be working on your process to show up. The second aspect is statistical in nature. You need enough samples to determine the characteristics of the distribution of sample means to find out if the process is stable, and, if stable, to make good estimates of the process mean and spread. Usually, 25 or more samples will be adequate as long as these samples contain at least 100 items. In other words, 25 samples of five tests each would be good.

## 3. Chart the Data.

As you gather data, enter it into the **NCSS** database. If your data come from an automated tool or from another data file, use the **NCSS** import capability to transfer data into the program.

## 4. Analyze the Data.

As you analyze the charts, it's probably best to begin with the range chart. Once you have found and corrected the assignable causes in the range chart, examine the Xbar chart in the same manner. You begin with the range chart because the range data are used to calculate the Xbar control limits. An out-of-control range point could cause the Xbar control limits to be too wide to detect out-of-control Xbar values.

The first place to look on the chart is for points outside the control limits. If the process has no assignable causes perturbing it, the sample means and ranges should vary by chance only. This means that the values on the control chart will form a fairly random pattern centered about the grand mean. A point would very rarely lie beyond the control limit unless something has changed in the process. As we mentioned previously, a point beyond the control limit is an almost certain indication of an assignable cause.

Next, look for trends/patterns. If the sample means and ranges vary by chance only, there should be no obvious pattern to the points on the control chart. On any given length of the chart, there should be about as many points above the centerline as below it and the line connecting the sample means should be fairly jagged as it moves from one point to another. Some nonrandom patterns that you might see include obvious trends, cycles, and clusters of points. Many of these unusual patterns can be detected through a series of tests called *runs tests*, discussed in detail earlier. Even if none of the points go beyond the control limits, the presence of any of these nonrandom patterns indicates that an assignable cause is changing the process.

## 5. Find and Correct Assignable Causes.

When you detect an assignable cause, analyze the process operation to find the cause. Correct the condition and prevent it from happening again. This problem-solving step is frequently very difficult and time consuming, but it is very important. The control chart itself can give you some very valuable clues as to when the problem began that may help you correlate the problem to a known change in the process. However, you will often have to draw on other tools, such as Pareto charts or cause and effect diagrams to find the problem. Don't overlook help from people involved in the process. They know the process better than anyone else and want to do the best job possible. They can be a great source of process improvements.

Once you have corrected a special cause, go back to the **NCSS** database and remove the subgroups that were affected by the special cause and rerun the analysis to calculate a new mean and control limits. Unless you know that a single value is in error, remove the entire sample, not just individual points that appear to be out of place. The importance of this step goes beyond just throwing away "bad data." If the points are not removed, they will cloud the rest of the analysis, making it impossible to get a good estimate of the amount of variation due to common causes. By removing the points, we will have a better "baseline" to use to detect future special causes when they occur.

# Chapter 251

# Attribute Charts

## Introduction

This procedure generates attribute control charts, including the p-, np-, c-, and u-charts. Attribute charts are useful for displaying count data such as the number of defectives or the number of defects on a certain item.

## Attribute Control Charts

Suppose we have a scatter plot with the number of defects on the vertical axis and time (such as hours, shifts, days, weeks, or months) on the horizontal axis. This scatter plot shows the nature of the number of defects over time. For example, we might see trends, shifts, sudden jumps, and so on. If we add horizontal limit lines to the plot to indicate standards, the scatter plot becomes a control chart. When the plots fall inside these limits lines, the process yielding the response is said to be in control. When the process yields responses that are outside these limits, the process is said to be out of control.

The limit lines set a range of “normal behavior.” They are based on past experience with the process and give a frame of reference for judging current outcomes. Because of natural variation in the process, the responses will not be exactly the same. They will bounce up and down. As long as the response stays within the limits, we need take no corrective action. However, once a measurement occurs outside the limits, we must investigate the cause and take appropriate corrective action.

Although there are many forms of control charts, they can be categorized as either variables or attributes control charts. Here, the term variable means that the data can take on any value. It does not have to be a whole number. A person’s weight or height are examples of variables data. (Variables charts are discussed in the chapter by that title.) Attributes, on the other hand, are things that can be counted, such as the number of defectives in a sample of 100 items or the number of scratches on a new car. It doesn’t make sense to talk about a half a scratch. The scratch either exists or it doesn’t. If you can put things in categories such as good or bad, acceptable or not acceptable, then they are attributes data.

### P-Chart

The charted value is the fraction defective in a sample of *n* items. Probability calculations are based on the binomial distribution.

### NP-Chart

The charted value is the number defective in a sample of *n* items per time period. The chart is very confusing if the sample size, *n*, is not constant from period to period. Probability calculations are based on the binomial distribution.

### C-Chart

The charted value is the number of defects on an item. Probability calculations are based on the Poisson distribution.

### U-Chart

The charted value is the average number of defects on *n* items. Probability calculations are based on the Poisson distribution.

## Formulas for Attribute Control Charts

Once we understand that a control chart is simply a plot of some measurement across time with appropriate limits shown as horizontal lines, the only question is how to determine these limits. The answer to this question depends on the situation. The following formulas give the mathematical details on how to set the limits for different types of measurements.

Suppose we have $k$ sets of samples. Let $n_i$ represent the sample size of the i$^{th}$ sample. Let $r_i$ represent the number of defectives (the number with the attribute of interest).

### P-Chart Calculations

The value plotted is

$$p_i = \frac{r_i}{n_i}$$

The center line is

$$\bar{p} = \frac{\sum_{i=1}^{k} r_i}{\sum_{i=1}^{k} n_i}$$

The control limits for the i$^{th}$ sample are

$$C.L. = \bar{p} \pm m\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$$

where $m$ is the multiplier, which is usually set to three.

### NP-Chart Calculations

The value plotted is $r_i$. The center line is

$$n_i \bar{p} = n_i \frac{\sum_{i=1}^{k} r_i}{\sum_{i=1}^{k} n_i}$$

The control limits for the i$^{th}$ sample are

$$C.L. = n_i \bar{p} \pm m\sqrt{n_i \bar{p}(1-\bar{p})}$$

where $m$ is the multiplier, which is usually set to three.

### C-Chart Calculations

The value plotted is $r_i$. The center line is

$$\bar{c} = \frac{\sum_{i=1}^{k} r_i}{k}$$

The control limits for the i[th] sample are

$$C.L. = \bar{c} \pm m\sqrt{\bar{c}}$$

where $m$ is the multiplier, which is usually set to three.

### U-Chart Calculations

In this chart $r_i$ represents the number of defects per item, and $n_i$ represents the number of items. The value plotted is

$$u_i = \frac{r_i}{n_i}$$

The center line is

$$\bar{u} = \frac{\sum_{i=1}^{k} r_i}{\sum_{i=1}^{k} n_i}$$

The control limits for the i[th] sample are

$$C.L. = \bar{u} \pm m\sqrt{\frac{\bar{u}}{n_i}}$$

## Runs Tests

Runs tests are discussed in the chapter on Variables Charts. Please turn to that chapter for further details.

## Data Structure

The data given below show how a set of attribute data is entered into **NCSS**. Each row gives the sample size (*Count*) and the number of defectives (*Rejects*). Twenty-seven rows of data are in the QATEST database. Only the first eight rows are shown here.

**QATEST dataset (subset)**

| Count | Reject |
|---|---|
| 100 | 1 |
| 99 | 2 |
| 100 | 2 |
| 100 | 1 |
| 100 | 1 |
| 100 | 0 |
| 100 | 2 |
| 100 | 0 |

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

## Variables Tab

This panel specifies the variables that will be used in the analysis.

### Variables

#### Sample Size Variable

This option specifies the variable containing the sample size, $n_i$. It is required for the P-, NP-, and U-charts.

#### Defects/Defectives Variable

This option specifies the variable containing the number of defectives for the P- and NP-charts or the number of defects for the C- and U-charts. It is the value of $r_i$.

#### Label Variable

An optional variable containing row labels that you may use to document your output. You can use dates (like Jan-23-95) as labels. First, enter your dates using the standard date format (like 06/20/93). In the Variable Info screen, change the format of the date variable to something like *mmm-dd-yyyy* or *mm-dd-yy*. The labels will be displayed as labels. Without changing the variable format, the dates will be displayed as long integer values.

### Specify Rows in Calculations

#### Specification Method

This option specifies how the rows that are used in the calculations are specified.

- **All Rows**
  All rows are used.

- **First Row - Last Row**
  The first and last row is specified.

- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.

- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

- **Row List**
  The rows used by in calculations are specified by the Row List box below.

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

### Row List

Specify sets of rows to be used in calculations. A separate set of calculations will be carried out for each set. Example (with three sets): 1-50, 75-150, 175-Last. Note that Specification Method must be set to Row List.

Rows that are not included in this list will still be plotted if they are included in the list of charted rows.

## Chart Type

### Chart Type

This option specifies the type of chart that is to be displayed. Possible charts are P, NP, C, and U.

## Specify Rows in Charts

### Specification Method

This option specifies how the rows that are used in the charts are specified.

- **All Rows**
  All rows are used.

- **First Row - Last Row**
  The first and last row is specified.

- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.

- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

## Select Chart Attributes

### Mean Line

Specifies whether to display a horizontal line representing the mean on the charts.

### Primary Control Limits

Specifies whether to display horizontal lines representing the primary control limits on the charts.

### Secondary Control Limits

Specifies whether to display horizontal lines representing the secondary control limits on the charts.

### Trend Line

Specifies whether to display a trend line on the charts.

### Runs

Specifies whether to add a label identifying those rows that failed a particular runs test.

### Row Labels

Specifies whether to label each row along the horizontal axis using the values in the Label Variable or the row number.

### Spec Limits

Specifies whether to display horizontal lines representing the specification control limits on the charts.

### Zones

Specifies whether to display horizontal lines representing the six horizontal zones.

### Use Runs Tests

Specifies whether to use the runs tests in determining out-of-control points.

### Label Out-of-Control Rows

Specifies whether to label rows that fall outside the control limits. If a Label Variable is used, the label specified there is used. Otherwise, the row number of the out-of-control point is given.

## Options Tab

This panel controls the calculation of the attribute chart.

### General Chart Options

#### Primary Multiplier

This option specifies the multiplier of sigma for the primary control limits. Usually, the famous 3-sigma limits are desired, so the multiplier is 3.

#### Secondary Sigma Multiplier

This option specifies the multiplier of sigma for the optional, secondary control limits. Usually, the secondary limits are ignored by setting this value to 0. Occasionally, a value of 2 is used.

#### Label Mean/CLs

Specifies whether to show the values of the mean and controls limits on the right of the chart. You can also add optional headings like LCL= or Mean= with the option “Yes-Labels.”

#### Robust Options

You can have **NCSS** scan your data and remove out-of-control rows from the calculations. Usually, you would perform this manually by repeatedly removing out-of-control rows. Occasionally you will want to obtain the final result without the manual intervention. These options define the automatic procedure that you want to use.

Caution: Accepted SPC procedure is to only remove out-of-control rows from the calculations if an assignable cause for the out-of-control situation can be found and corrected. If a cause cannot be found and corrected, you should not remove the out-of-control row. Hence, you should be careful about using this automated procedure, since it does not require you to find assignable causes for the out-of-control points.

#### Robust Iterations

This option specifies the number of robust iterations (cycles through the data looking for out-of-control subgroups) that you want. Usually, one or two should be sufficient.

If you want to skip the robust estimation entirely, enter a 0 here.

#### Robust Multiplier

This option specifies the control limit multiplier to be used during the robust estimation. Usually you would enter a 3 here. Occasionally you might want to adjust this value slightly.

#### Rows Skipped

If you print out individual row labels along the horizontal axis and you have many rows, the labels may overwrite each other. This option lets you skip *x* number of labels. For example, if you only wanted to display every other label, you would enter a 1 here. If you only wanted to display every fifth label, you would enter a 4 here.

## Xbar Chart Options

### Fixed P or C Type

A fixed value for P (or C) can be input as a constant or as the first row in a specified variable. This option specifies where to find the fixed value. It specifies which of the following two option boxes contain the fixed mean value.

### P or C Constant

This option lets you specify a fixed value for P (or C). It requires the appropriate selection of *Constant* in the last option.

### P or C Variable

This option lets you specify a fixed value for P (or C). To use this, you would enter the fixed value as the first value in this variable on the data base. This might be convenient when several databases (each with different fixed values) are being run. It requires the appropriate selection of Variable in the above option.

## Specification Limits

### Lower Specification Limit

This option lets you specify the optional lower specification limit for display on your charts.

### Upper Specification Limit

This option lets you specify the optional upper specification limit for display on your charts.

# Reports Tab

The following options control the format of the reports.

## Specify Reports

### Chart Summary Section

This option controls the display of this report.

### Exception Section

This option controls the display of this report.

## Specify Charts

### Attribute Chart

This option controls the display of the attribute chart.

## Report Options

### Precision

This allows you to specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some

numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

### Page Title

This option specifies a title to appear at the top of each page.

### Plot Subtitle

This option specifies a subtitle to appear at the top of each plot.

## Charts Tab

This panel sets the options used to define the appearance of the chart.

### Vertical and Horizontal Axis

#### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

#### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

### Attribute Chart Settings

#### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

## Symbols-Lines Tab

This panel specifies the plotting symbols and lines used on the charts.

### Symbols

#### Out of Control

Specify the symbol used to display the out-of-control points. Click the button on the right to display the symbol modification window.

#### In Control

Specify the symbol used to display the in-control points. Click the button on the right to display the symbol modification window.

### Lines

#### Connecting Line - Zone Lines

These options specify the color, width, and style of the various lines that make up the control chart.

## Storage Tab

This panel controls the automatic storage of the proportions (or average defects) on the current database.

### Storage Variable

#### Store P Values in Variable

You can automatically store the proportion defective (or average defects) of each row into the variable specified here. Warning: Any data already in this variable is replaced. Be careful not to specify variables that contain important data.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Creating an Attribute Chart

This section presents an example of how to generate a P-chart. The data used are found in the QATEST database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Attribute Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Attribute Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Attribute Charts**. The Attribute Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the Attribute Charts window, select the **Variables tab**.
- Double-click in the **Sample Size Variable** text box. This will bring up the variable selection window.
- Select **COUNT** from the list of variables and then click **Ok**. "COUNT" will appear in the Sample Size Variables box.
- Double-click in the **Defects/Defectives Variable** text box. This will bring up the variable selection window.
- Select **REJECT** from the list of variables and then click **Ok**. "REJECT" will appear in the Defects/Defectives Variable box.

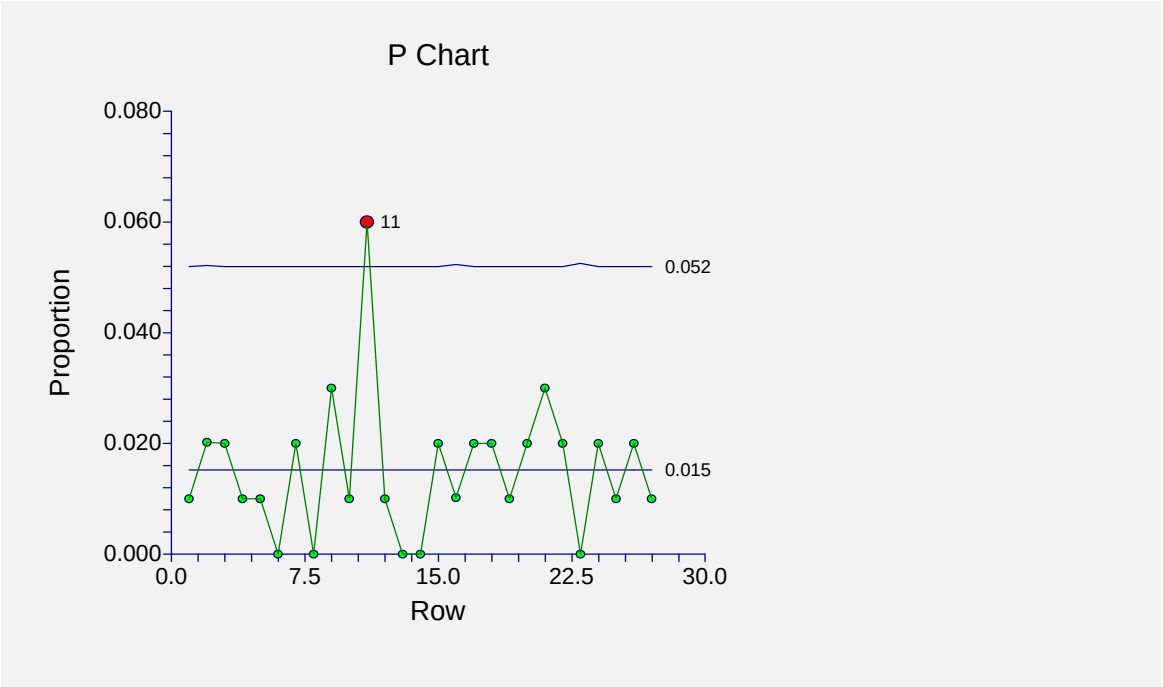
**4 Specify the chart.**

- On the Attribute Charts window, select the **Charts tab**.
- Enter **0** in the **Minimum** text box.
- Press the **Tick Label Settings…** button under Vertical Axis. This will bring up the Settings of Tick Label Settings window.
- Select **3** in the **Decimals** list box.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## P-Chart




This plot displays a P-chart that shows the fraction defective over time.

## Control Limits Section

**Control Limits Section**

| Control Limit | Proportion |
|---|---|
| Lower | 0.000000 |
| Upper | 0.052510 |

This section displays the values of the lower and upper control limits.

## Estimation Summary Section

**Estimation Summary Section**

| Type | Variables | Total | Mean |
|---|---|---|---|
| Samples | COUNT | 2694 | 99.777778 |
| Defects | REJECT | 41 | 1.518519 |

| Type | Proportion |
|---|---|
| User Specified | 0 |
| Data* | 0.015219 |
| Robust | 0.015219 |
| Number of Rows | 27 |

This report gives the numerical details of the analysis. We'll now define each of the numbers appearing on the report.

### Samples

This line gives the sample size variable's name, the total, and the average.

### Defects

This line gives the defects variable's name, the total, and the average.

### Proportion

This column gives the estimated proportion for each estimation method.

**User Specified** This is used if a fixed proportion was specified by the user.

**Data** This is the estimate of the proportion defective that is calculated from the data. The star (*) by the word "Data" indicates that this is the estimate that was used in the current chart.

**Robust** This is the estimated proportion when the robust estimation method is used.

## Out-of-Control List

**Out-of-Control List**

| Row | Proportion | Row Label | Reason |
|---|---|---|---|
| 11 | .060000 | 11 | beyond control limits |

This report provides a list of the rows that failed one of the runs tests (including being outside the control limits). The Reason column names the particular runs test that was failed.

# Chapter 252

# Levey-Jennings Charts

## Introduction

This procedure generates Levey-Jennings control charts on single variables. It finds out-of-control points using the Westgard rules.

## Levey-Jennings Control Charts

The Levey-Jennings control chart is a special case of the common Shewart Xbar (variables) chart in which there is only a single stream of data and sigma is estimated using the standard deviation of those data. The formula for the standard deviation *s* is

$$s = \sqrt{\frac{\sum_{k=1}^{n}\left(x_k - \bar{x}\right)^2}{n-1}}$$

where the mean is estimated using

$$\bar{x} = \frac{\sum_{k=1}^{n} x_k}{n}$$

Control limits are

$$\left(L_{low}, L_{high}\right) = \bar{x} \mp ms$$

where *m* is usually 1, 2, or 3.

## Westgard Rules

Individual values are tested to determine if they are in, or out, of control using a set of five rules called the Westgard rules after their originator. They are specified in Westgard *et al.* (1981). These rules indicate which rows in a variable (column of numbers) are ‘out-of-control’. When any of these rules is violated, the process behind the numbers is ‘out-of-control’ and should be stopped and investigated.

The Westgard Rules are

**1S3**: One value beyond 3*sigma from the mean.

**2S2**: Two consecutive values either greater than, or less than, 2*sigma from the mean.

**RS4**: A difference between consecutive values greater than 4*sigma.

**4S1**: Four consecutive values greater than, or less than, 1*sigma from the mean.

**10X**: Ten consecutive values all greater than, or less than, the mean.

# Data Structure

The data are entered in a single variable (column) of the spreadsheet. As an example, you can look at the WESTGARD.S0 database. Often, variables are entered as pairs, but this is not necessary.

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

## Variables Tab

This panel specifies the variables that will be used in the analysis.

### Variables

#### Data Variables

These are the variables to be analyzed. A separate chart is generated for the values in each variable. Note that the rows represent the way the data were received through time. That is, row one gives the first value obtained, row two gives the second value, and so on.

#### Label Variable

An optional variable containing row labels for the horizontal axis of the chart.

You can use dates (like Jan-23-95) as labels. First, enter your dates using the standard date format (like 06/20/93). In the Variable Info screen, change the format of the date variable to something like *mmm-dd-yyyy* or *mm-dd-yy*. The labels will be displayed as labels. Without changing the variable format, the dates will be displayed as long integer values.

### Specify Rows in Calculations

#### Specification Method

This option specifies how the rows that are used in the calculations are specified.

- **All Rows**
  All rows are used.

- **First Row - Last Row**
  The first and last row is specified.

- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.

- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

- **Row List**
  The rows used by in calculations are specified by the Row List box below.

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

### Row List

Specify sets of rows to be used in calculations. A separate set of calculations will be carried out for each set. Example (with three sets): 1-50, 75-150, 175-Last. Note that Specification Method must be set to Row List.

Rows that are not included in this list will still be plotted if they are included in the list of charted rows.

## Specify Rows in Charts

### Specification Method

This option specifies how the rows that are used in the charts are specified.

- **All Rows**
  All rows are used.

- **First Row - Last Row**
  The first and last row is specified.

- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.

- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Specification Method is set to First Row - Last Row.

### N

This option designates the value of N. This option is only used when Specification Method is set to First N Rows or Last N Rows.

## Select Chart Attributes

### Mean Line

Specifies whether to display a horizontal line representing the mean.

### Sigma Limits

Specifies whether to display horizontal lines representing the control limits for each multiple.

### Trend Line

Specifies whether to display a trend line on the charts.

### Row Labels

Specifies whether to label each row along the horizontal axis using the values in the Label Variable or the row number.

### Spec Limits

Specifies whether to display horizontal lines representing the specification limits.

### Use Westgard Rule

Specify whether to include rows that violate this rule on the plot and in the exceptions report.

The codes are:

- **1S3**
  1 value beyond 3*sigma from the mean.
- **2S2**
  2 consecutive values >, or <, 2*sigma from the mean.
- **RS4**
  A difference between consecutive values > 4*sigma.
- **4S1**
  4 consecutive values >, or <, 1*sigma from the mean.
- **10X**
  10 consecutive values >, or <, the mean.

# Options Tab

These options determine the type of chart that you want displayed.

## General Chart Options

### 1-Sigma, 2-Sigma, and 3-Sigma Multipliers

This option specifies the multiplier of sigma for each set of control limits. Usually, the multipliers are set to 1, 2, and 3.

### Label Mean and Control Limits

Specifies whether to show the values of the mean (center) and controls limits on the right of the chart. You can also add optional headings like LCL= or Mean= with the option "Yes-Labels."

### Rows Skipped in Labels

If you print out individual row labels along the horizontal axis and you have many rows, the labels may over-write each other. This option lets you skip *x* number of labels. For example, if you only wanted to display every other label, you would enter a 1 here. If you only wanted to display every fifth label, you would enter a 4 here.

## Mean Options

### Mean From

This option specifies how the mean is determined. Usually, it is calculated from the data. But occasionally, a fixed value is used. Select Data to calculate the mean from the data, Constant to use the value in the Mean Constant box, or Variable to read the mean from a specific variable on the database.

### Mean Constant

This value is used as the value of the mean when Mean From is set to Constant.

### Mean Variable

Values in the rows of this variable (column) are used as the value of the means when Mean From is set to Variable.

Note that the value in row one is used for the variable in column 1 of the spreadsheet, the value in row two is used for the variable in column 2, and so on. If you have selected variables number 10 and 15 as your Data Variables, then rows 10 and 15 will contain the values of the fixed values of the means of these variables.

## Sigma Options

### Sigma From

This option specifies how sigma is determined. Usually, it is calculated from the data. But occasionally, a fixed value is used. Select Data to calculate sigma from the data, Constant to use the value in the Sigma Constant box, or Variable to read sigma from a specific variable on the database.

### Sigma Constant

This value is used as the value of sigma when Sigma From is set to Constant.

### Sigma Variable

Values in the rows of this variable (column) are used as the value of the sigma when Sigma From is set to Variable.

Note that the value in row one is used for the variable in column 1 of the spreadsheet, the value in row two is used for the variable in column 2, and so on. If you have selected variables number 10 and 15 as your Data Variables, then rows 10 and 15 will contain the values of the fixed values of sigma of these variables.

## Specification Limits

### Lower and Upper Spec Limit

These options specify specification limits for display on the Levey-Jennings chart.

### Target Spec

This option specifies an optional target specification for display on the Levey-Jennings chart.

# Reports Tab

The following options control the format of the reports.

## Specify Reports

### Numeric Reports – Out-of-Control List

Each of these options control the display of the corresponding report.

## Specify Charts

### Levey-Jennings Chart

This option controls the display of the Levey-Jennings chart.

## Report Options

### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Variable Names

This option lets you select whether to display variable names, variable labels, or both.

### Decimal Places

Set the number of decimal places displayed on the reports. For example, selected 2 here instructs the program to display the value 1.2362142 as 1.24.

*Single* displays an unformatted, seven-digit number. *Double* displays an unformatted, fourteen-digit number.

### Page Title

This option specifies a title to appear at the top of each page.

## Levey-Jennings Charts Tab

This panel sets the options used to define the appearance of the Levey-Jennings control chart.

### Vertical and Horizontal Axis

#### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

#### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

#### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

#### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

### Attribute Chart Settings

#### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

#### Top Title 2

Enter text to appear as the second title line at the top of the plot. Enter {M} to cause the mean, standard deviation, and coefficient of variation to be displayed.

### Bottom Title 2

Enter text to appear as the second title line at the bottom of the plot. Enter {M} to cause the mean, standard deviation, and coefficient of variation to be displayed.

## Symbols - Lines Tab

This panel specifies the plotting symbols and lines used on the charts.

### Symbols

#### Out of Control

Specify the symbol used to display the out-of-control points. Click the button on the right to display the symbol modification window.

#### In Control

Specify the symbol used to display the in-control points. Click the button on the right to display the symbol modification window.

### Lines

#### 1-Sigma Lines - Specification Lines

These options specify the color, width, and style of the various lines that make up the control chart.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

#### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Creating a Levey-Jennings Control Chart

This section presents an example of how to generate a Levey-Jennings control chart. The data are found in the WESTGARD database. We will analyze the variable Test3 on this database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Levey-Jennings Charts window.

**1 Open the WESTGARD dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **WESTGARD.S0**.
- Click **Open**.

**2 Open the Levey-Jennings Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Levey-Jennings Charts**. The Levey-Jennings Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

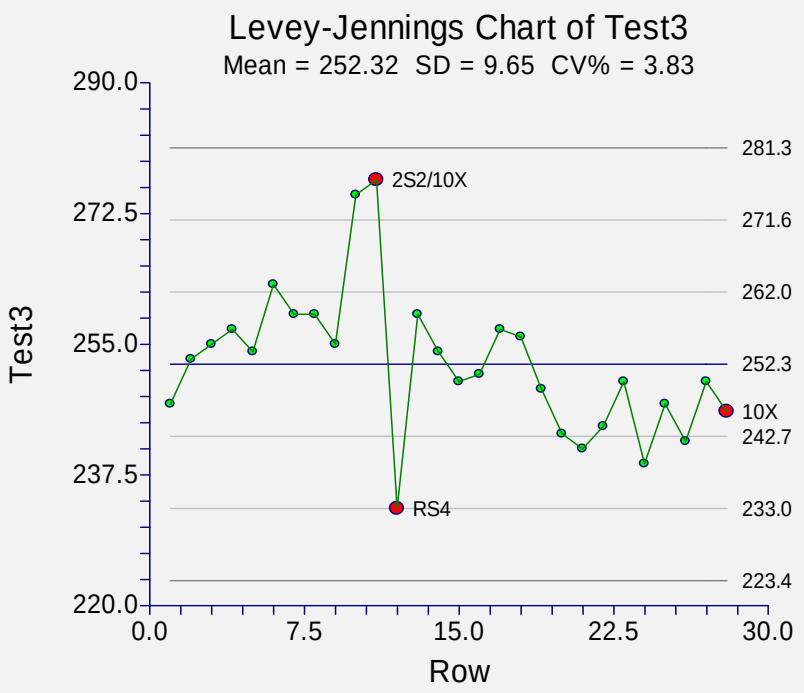
**3 Specify the variables.**

- On the Levey-Jennings Charts window, select the **Variables tab**.
- Set the **Data Variables** box to **Test3**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Levey-Jennings Control Chart




This plot displays the Levey-Jennings control chart. The overall mean (center-line) and three sets of control limits are shown. Notice that three rows are out of control. The next report gives the numerical details of the charts and lists those rows that failed at least one of the control tests.

## Numerical Reports

**Descriptive Statistics Section for Test3**

| Rows Used in Calculations | Mean | SD | CV% | Row Count |
|---|---|---|---|---|
| 1-28 | 252.32 | 9.65 | 3.83 | 28 |

**Control Limits Section for Test3**

| Rows Used in Calculations | Mean | Lower 3-Sigma | Upper 3-Sigma | Lower 2-Sigma | Upper 2-Sigma |
|---|---|---|---|---|---|
| 1-28 | 252.32 | 223.36 | 281.28 | 233.01 | 271.63 |

**Out-of-Control List for Test3**

| Row | Value | Reason |
|---|---|---|
| 11 | 277 | 2S2: 2 consecutive values >, or <, 2 sigma |
| | | 10X: 10 consecutive values >, or <, mean |
| 12 | 233 | RS4: consecutive difference > 4 sigma |
| 28 | 246 | 10X: 10 consecutive values >, or <, mean |

The Descriptive Statistic section displays the values of the calculated mean, standard deviation, and coefficient of variation (which is expressed as a percentage). The Control Limits section displays the 2-sigma and 3-sigma control limits. The Out-of-Control List gives a list of all rows that failed at least one of the Westgard rules.

# Chapter 253

# Pareto Charts

## Introduction

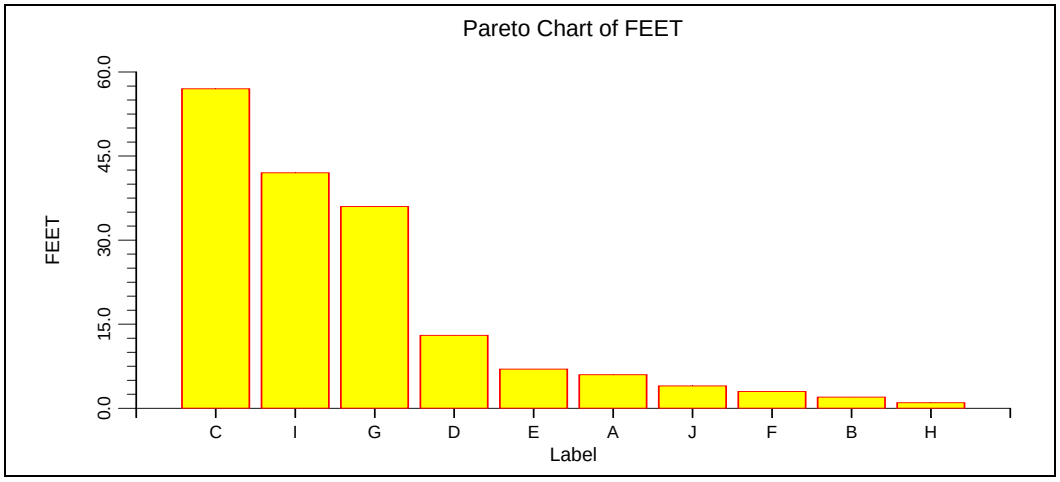
An Italian economist, Vilfredo Pareto (1848-1923), noticed a great inequality in the distribution of wealth. A few people owned most of the wealth. J. M. Juran found that this same phenomenon of the "vital few and the trivial many" applied to many areas of SPC. He is credited with coining the terms "Pareto chart" and "Pareto analysis" to represent this phenomenon.

In quality control, Pareto analysis refers to the tendency for the bulk of the quality problems to be due to a few of the possible sources. Hence, by isolating and correcting the major problem areas, you obtain the greatest increase in quality. The Pareto chart is a graphic display that emphasizes the Pareto principle using a bar graph in which the bars are arranged in decreasing magnitude.

**NCSS** provides two Pareto chart styles as well as a numerical report.

## Pareto Charts

The following plot shows a Pareto chart depicting the number of defective board-feet (in 100's) from ten different mills. Notice that three mills account for almost 80% of the defective product. Obviously, efforts should be concentrated on correcting defects in these three mills.



# Data Structure

The table below shows the data for the above Pareto chart. It gives the number of defective board-feet (in 100's) from ten different mills (labeled A - J). These data are contained on the QATEST database.

**QATEST dataset (subset)**

| Label | Feet |
|---|---|
| A | 6 |
| B | 2 |
| C | 57 |
| D | 13 |
| E | 7 |
| F | 3 |
| G | 36 |
| H | 1 |
| I | 42 |
| J | 4 |

# Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

## Variables Tab

This panel specifies the variables that will be used in the analysis.

### Data and Label Variables

#### Data Variables

This (required) option specifies which variables on the database contain the actual data values. If more than one variable is specified, the number of charts generated depends on the status of the Data Item option (described below).

Note that all data must be positive for the Pareto Chart. Negative values are ignored.

#### Chart Arrangement

This option specifies the way in which the data are to be arranged on the Pareto chart.

- **Each Row**

  This option causes a separate chart to be generated for each of the Data Variables specified (see above). Labels may be set using the Label Variable option (see above). Each row of data becomes a bar on the Pareto chart. Note that the Category Variable is ignored when this option is used.

- **Total By Variable**

  This option causes one chart to be constructed using all of the Data Variables specified (see above). The total for each variable becomes a bar on the Pareto chart. Note that the Category Variable and the Label Variable are ignored when this option is used.

- **Average By Variable**

  This option causes one chart to be constructed using all of the Data Variables specified (see above). The average for each variable becomes a bar on the Pareto chart. Note that the Category Variable and the Label Variable are ignored when this option is used.

- **Total By Category**

  This option causes one chart to be constructed for each of the Data Variables specified (see above). The average of the Data Variable for each unique value of the Category Variable becomes a bar on the Pareto chart. Note that the Label Variable is ignored when this option is used.

- **Average By Category**

  This option causes one chart to be constructed for each of the Data Variables specified (see above). The average of the Data Variable for each unique value of the Category Variable becomes a bar on the Pareto chart. Note that the Label Variable is ignored when this option is used.

### Label Variable

An optional variable containing labels for the individual data values may be entered here. Note that this option is only used when a single variable is analyzed. You can use dates (like Jan-23-95) as labels. Here is how. First, enter your dates using the standard date format (like 06/20/93). In the Variable Info screen, change the format of the date variable to something like *mmm-dd-yyyy* or *mm-dd-yy*. The labels will be displayed as labels. Without changing the variable format, the dates will be displayed as long integer values.

## Category Specification

### Category Variable

An optional categorical (grouping) variable may be specified. If it is used, the Data Variable variable will be summed (or averaged) by the values of this variable. Hence, this must be a discrete variable.

### Minimum Value

Values on the Pareto chart less than or equal to this value are lumped together into one category. This combined category is labeled using the Other Category Label.

### Below Minimum Category Label

This option specifies the label to be displayed for the combined value when the MinimumValue is used.

## Specify Rows

### Row Specification Method

This option specifies how the rows that are used in the calculations and displayed on the charts are specified.

- **All Rows**
  All rows are used.
- **First Row - Last Row**
  The first and last row is specified.
- **First N Rows**
  The first N rows on the dataset are used. The value of N is specified below.
- **Last N Rows**
  The last N rows on the dataset are used. The value of N is specified below.

## Specify Rows – First Row / Last Row Details

### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Row Specification Method is set to First Row - Last Row.

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Row Specification Method is set to First Row - Last Row.

## Specify Rows – N Details

### N

This option designates the value of N. This option is only used when Row Specification Method is set to First N Rows or Last N Rows.

## Pareto Chart Bars

### Label Position

This option specifies if and where the percentage value should be displayed.

## Pareto Chart Bars - Fill

### Color

This option specifies the color of the interior portion of the bars.

### Pattern

This option specifies the pattern of the interior portion of the bars.

## Pareto Chart Bars - Outline

### Color

This option specifies the color of the edge of the bars.

### Width

This option specifies the width of the edge of the bars.

## Pareto Chart Bars - Width

### Select Bar Width Parameter

This option lets you designate whether to specify the bar width using the actual amount or the percent of space between the bars.

### Amount

When the Select Bar Width Parameter is set to Amount, the option gives the width of the bars.

### Percent Empty Space

When the Select Bar Width Parameter is set to Percent Empty Space, the option gives the percent of the total space that is to be between the bars.

## Cumulative Line

### Symbol

This option specifies the color, size, and type of plotting symbol used in the cumulative Pareto chart.

### Line Width

This option specifies the width of the cumulative line.

# Reports Tab

The options on this panel control the format of the reports.

## Select Charts

### Regular Chart

This option specifies whether to display the regular pareto chart.

### Cumulative Chart

This option specifies whether to display the cumulative pareto chart.

## Select Reports

### Numeric Report

This option specifies whether to display the numeric report.

## Report Options

### Precision

This option specifies the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

### Variable Names

This option lets you select whether to display variable names, variable labels, or both.

## Page Title

### Page Title

This option specifies a title to appear at the top of each page.

# Pareto Chart Tab

This panel sets the options used to define the appearance of the chart.

## Vertical and Horizontal Axes

### Label (Y and X)

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Maximum (Y)

This option specifies the maximum value displayed on the vertical (Y) axis. If left blank, it is calculated from the data. Note that the minimum value is always set to zero.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the vertical axis.

### Show Grid Lines

This check box indicates whether the grid lines that emanate from the vertical axis should be displayed.

### Tick Label Settings

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

## Pareto Chart Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Titles

#### Plot Title

This is the text of the title. The characters *{Y}*, *{X}*, and *{G}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

#### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

## Example 1 – Creating a Pareto Chart

This section presents an example of how to generate a Pareto chart. The data used are shown in the table at the beginning of the chapter and are found in the QATEST database.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Pareto Charts window.

**1 Open the QATEST dataset.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

**2 Open the Pareto Charts window.**

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Pareto Charts**. The Pareto Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

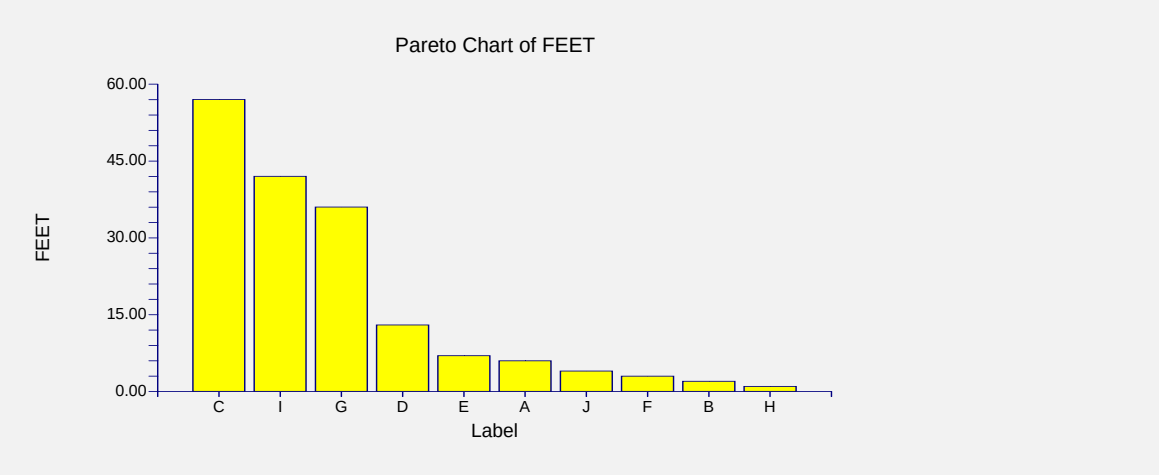
- On the Pareto Charts window, select the **Variables tab**.

- Double-click in the Data Variables text box. This will bring up the variable selection window.
- Select **FEET** from the list of variables and then click **Ok**. "FEET" will appear in the Data Variables box.
- Double-click in the Row Label Variable text box. This will bring up the variable selection window.
- Select **Label** from the list of variables and then click **Ok**. "Label" will appear in the Label Variable box.

**4 Run the procedure.**

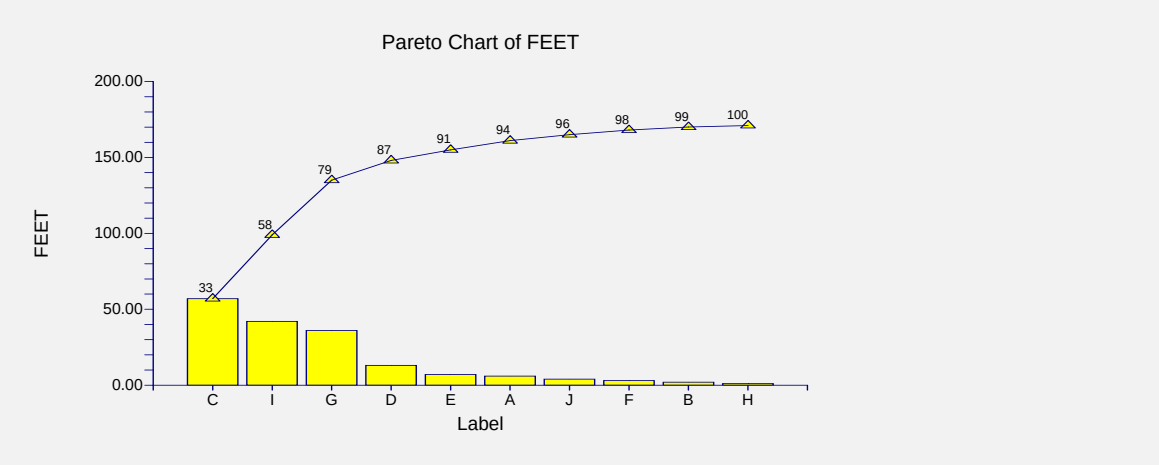
- From the Run menu, select Run Procedure. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Pareto Chart




This plot displays the typical descending bar chart. Note that the scale on the left is in terms of the individual items (mills).

## Cumulative Pareto Chart




This section displays the cumulative Pareto chart. Note that this version of the Pareto chart combines the bar chart version with a line representing the cumulative total for each bar. The

cumulative percentage is displayed above the plotting symbol. For example, we see that 79% of the defects are caused by mills C, I, and G.

## Pareto Numeric Report

**Pareto Numeric Report**

| Label | FEET | Cumulative FEET | Percent | Cumulative Percent |
|---|---|---|---|---|
| A | 6 | 6 | 3.51 | 3.51 |
| B | 2 | 8 | 1.17 | 4.68 |
| C | 57 | 65 | 33.33 | 38.01 |
| D | 13 | 78 | 7.60 | 45.61 |
| E | 7 | 85 | 4.09 | 49.71 |
| F | 3 | 88 | 1.75 | 51.46 |
| G | 36 | 124 | 21.05 | 72.51 |
| H | 1 | 125 | .58 | 73.10 |
| I | 42 | 167 | 24.56 | 97.66 |
| J | 4 | 171 | 2.34 | 100.00 |

This report gives the numerical details of the analysis.

# Example 2 – Using Several Variables

This section presents an example of how to generate a Pareto chart of the total for several variables. The data used are the values of S1 - S5 found in the QATEST database. Suppose, for the moment, that these five variables represent the fifty daily numbers of defects produced by each of five shifts.

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Pareto Charts window.

### 1 Open the QATEST dataset.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **QATEST.s0**.
- Click **Open**.

### 2 Open the Pareto Charts window.

- On the menus, select **Graphics**, then **Quality Control Charts**, then **Pareto Charts**. The Pareto Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the variables.

- On the Pareto Charts window, select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** through **S5** from the list of variables and then click **Ok**. “S1-S5” will appear in the Data Variables box.
- Select **Total By Variable** from the **Chart Arrangement** list box.

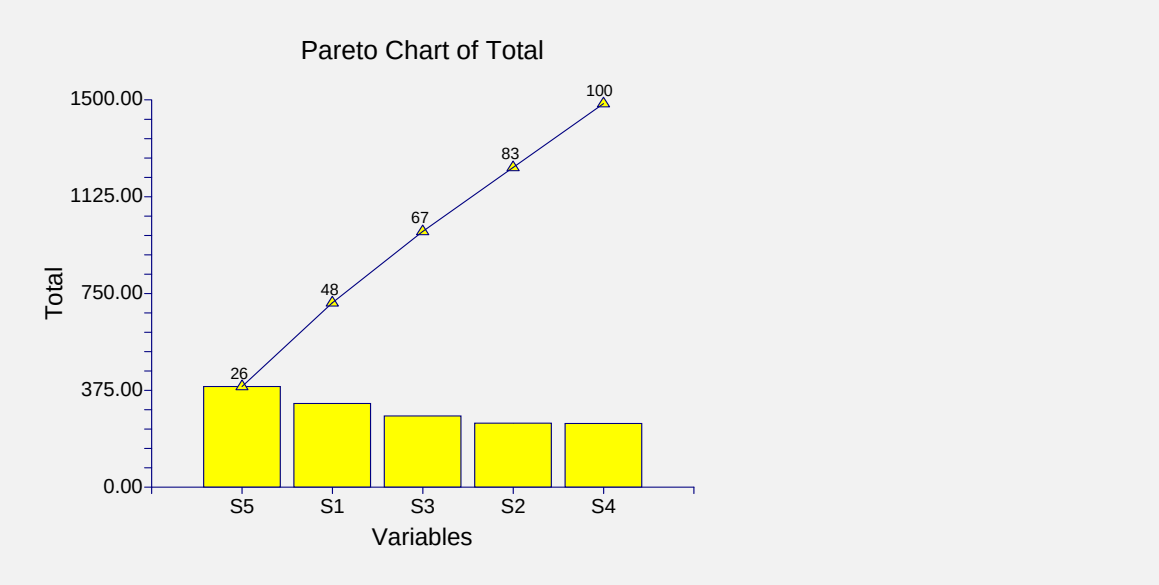
**4 Specify the reports.**

- On the Pareto Charts window, select the **Reports tab**.
- Click **Regular Chart** so that it is **not checked**.
- Click **Numeric Report** so that it is **not checked**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Pareto Chart




Notice that a somewhat uniform defect rate in the five shifts is reflected here by almost identical bars and by an almost straight cumulative line.

# Chapter 254

# R & R Study

## Introduction

A repeatability and reproducibility (R & R) study (sometimes called a gauge study) is conducted to determine if a particular measurement procedure is adequate. If the measurement variation is small relative to the actual process variation, the measurement procedure is adequate. If it is not, the measurement procedure must be improved before it can satisfactorily monitor the process. For example, if your manufacturing specifications are in millimeters, but your measuring device provides readings only in centimeters, you are trouble.

R & R studies separate process variation into that due to the measurement procedure and that due to the production process itself. The measurement variation is further divided into that due to the appraiser (reproducibility) and that due to the measuring device (repeatability).

It is important to emphasize that an R & R study is concerned with the precision of the measurement process. Data for R & R studies come from experiments especially designed for that purpose and that purpose only! Do not attempt to combine these studies with other experiments that you are conducting.

Several booklets are available that discuss R & R studies in detail. We recommend Barrentine (1991) and AIAG (1995). Although both of these concentrate on the “control chart” approach, they mention the analysis of variance approach—which we use here. The AIAG booklet states that the control chart approach is to be used only when software to analyze your data with the analysis of variance approach is not available.

## Data Structure

Burdick and Larsen (1997) discuss an R & R study conducted to determine the capability of a procedure for monitoring the chemical content of a large tank. Ten samples are taken from the tank. A random sample of three operators is selected for the study. Each operator measures the chemical content of each of the ten samples three times using the same measurement device. The operator’s measurements are made in random order. It is assumed that the operators are experienced so that no learning occurs during the study. The ninety values of acid concentration are recorded in the RRSTUDY dataset and displayed in the following table. Note that the results of a particular trial (a measurement by each of the three operators) are recorded a single row. Since each sample is measured three times by each operator, the results for each sample use three rows of the dataset.

An alternate way of entering these data is given at the end of this chapter.

**RRSTUDY dataset**

| Sample | Op1 | Op2 | Op3 |
|---|---|---|---|
| 1 | 67 | 66 | 69 |
| 1 | 68 | 68 | 67 |
| 1 | 68 | 68 | 68 |
| 2 | 67 | 67 | 67 |
| 2 | 66 | 67 | 66 |
| 2 | 66 | 68 | 66 |
| 3 | 68 | 70 | 68 |
| 3 | 68 | 70 | 68 |
| 3 | 67 | 68 | 68 |
| 4 | 67 | 70 | 67 |
| 4 | 67 | 68 | 68 |
| 4 | 67 | 70 | 68 |
| 5 | 68 | 70 | 69 |
| 5 | 68 | 70 | 68 |
| 5 | 68 | 70 | 69 |
| 6 | 69 | 71 | 70 |
| 6 | 68 | 70 | 70 |
| 6 | 69 | 70 | 70 |
| 7 | 67 | 68 | 68 |
| 7 | 67 | 68 | 68 |
| 7 | 67 | 68 | 69 |
| 8 | 75 | 75 | 75 |
| 8 | 74 | 75 | 75 |
| 8 | 74 | 75 | 75 |
| 9 | 67 | 69 | 68 |
| 9 | 67 | 68 | 68 |
| 9 | 67 | 69 | 68 |
| 10 | 66 | 68 | 66 |
| 10 | 66 | 66 | 66 |
| 10 | 66 | 66 | 66 |

## Missing Values

Missing values are not allowed in this analysis. The confidence limits are based on formulas for experiments in which no data are missing. If you have missing values, you should resolve them by removing the measurements for the sample with the missing data from the analysis. The bottom line is this—make sure you do not allow missing values!

# The Analysis of Variance Approach to R & R

The analysis of variance model of this experimental design is

$$Y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + E_{ijk}$$

where $i$=1,…,I; $j$=1,…,$J$; $k$=1,…,$K$; $I$=10; $J$=3; $K$=3; and $P_i$, $O_j$, $(PO)_{ij}$, $E_{ijk}$ are jointly independent normal random variables with means of zero and variances $\sigma_P^2$, $\sigma_O^2$, $\sigma_{PO}^2$, and $\sigma_E^2$, respectively. These variances are often referred to as *variance components*. We let $S$ represent the samples (parts), $O$ represent the operators (appraisers), and $E$ represent the random error.

In terms of this model, repeatability is $\sigma_E^2$, reproducibility is $\gamma_1 = \sigma_O^2 + \sigma_{PO}^2$, and the total variability associated with the measurement procedure is $\gamma_2 = \sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2$, which may be called the R & R value. The process (sample-to-sample) variability is represented by $\sigma_P^2$. A ratio that compares process variability to measurement variability is

$$\delta = \frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}$$

Several indices have been devised to summarize the results of such an R & R study. Many are based on the above quantities. For example, the automotive group defines the signal-to-noise ratio as

$$SNR = \sqrt{\delta} = \sqrt{\frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}}$$

and the number of distinct product categories that can be reliably distinguished by the measurement procedure as

$$Distinct\ Categories = \sqrt{2\delta} = \sqrt{\frac{2\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}}$$

Two popular measures that compare the measurement variance to the tolerance, where tolerance is the difference between the upper specification limit (USL) and lower specification limit (LSL), are the *measurement error ratio*

$$M = \frac{3\sqrt{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}}{USL - LSL} \times 100\%$$

and the *precision-to-tolerance ratio* (P/T)

$$PT = \frac{6\sqrt{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}}{USL - LSL} \times 100\%$$

All of these quantities are estimated in the analysis of variance approach with confidence intervals as well as point estimates.

The goal of the analysis is to estimate these quantities and determine if they fall within previously set guidelines.

# Procedure Options

This section describes the options available in this procedure.

## Variables Tab

This panel specifies the variables used in the analysis.

### Variables

#### Sample (Part) Variable

The values in this variable identify which sample (part) is represented on each row. The values may be numbers or text.

#### Appraiser (Operator) Variable

This variable is optional and is only used when a single Measurement Variable is specified. When this option is blank, you must specify at least two Measurement Variables.

The values in this variable identify which appraiser (operator) is represent on each row. The values may be numbers or text.

#### Measurement Variable(s)

The one or more variables specified here contain the measurements (scores). When only one measurement variable is specified, you must specify an Appraiser Variable. When multiple measurement variables are specified, you must leave the Appraiser Variable blank.

When more than one variable is specified, each variable contains the results for a particular appraiser. Each row represents the measurements of a part or sample on one trial. If you have multiple trials, you will have multiple rows.

### Specification Limits

#### Lower Spec Limit

This optional value is the lower specification limit. These limits are not control limits but the actual specification limits set by the manufacturer. They are used by the program to determine the tolerance, which is calculated using the formula: tolerance = Upper Spec Limit - Lower Spec Limit. It is not necessary to enter this value if you do not want to calculate statistics that involve the tolerance.

#### Upper Spec Limit

This optional value is the upper specification limit. These limits are not control limits but the actual specification limits set by the manufacturer. They are used by the program to determine the tolerance, which is calculated using the formula: tolerance = Upper Spec Limit - Lower Spec Limit. It is not necessary to enter this value if you do not want to calculate statistics that involve the tolerance.

#### Target Spec

This optional value specifies the target value of the item being studied. This value is used to calculate the deviation from target in the Means Report. It may be omitted.

## Sigma Multiplier

### Sigma Multiplier

The multiplier of the standard deviation that defines the percent of the normal distribution that is compared to the tolerance or the process variability. This value establishes the magnitude of the range of the measurement variable.

The most common value used is 5.15. This value is used because the mean plus or minus (5.15)/2 sigma contains 99.0% of the area under the normal distribution curve. Other popular choices are 6.00 sigma which contains 99.7% and 4.00 sigma which contains 95.0%.

# Reports Tab

The following options control which plots and reports are displayed.

## Specify Reports

### EMS Report – Means Report

Specify whether to display the indicated report.

## Specify Plots

### Means Plots – Residual Plots

Specify whether to display the indicated plots.

## Plot Legend

### Show Legend

Specifies whether to display the legend.

### Legend Text

Specifies legend label. A {G} is replaced by the word ‘Appraiser.’

## Report Options

### Confidence Level

The value of confidence coefficient (in percentage terms) for the confidence intervals. Usually, this number will range from 90 to 99.9. A common choice for confidence limits of variance components is 90. You should determine a value appropriate for your particular study.

### Precision

Specify the precision of numbers in the report. Single precision will display seven-place accuracy, whereas the double precision will display thirteen-place accuracy.

### Variable Names

Indicate whether to display the variable names or the variable labels.

### Value Labels

Indicate whether to display the data values or their labels.

### Appraiser Label

This options specifies the phrase used in the output reports to represent the appraisers or operators. This option is only used when several Measurement Variables are specified.

## Decimal Places

### Percents - Variances

These options let you specify the number of decimal places displayed in the reports. Select 'General' if you want to see the most digits possible. Your selection here does not change the precision of the calculations. All calculations use double precision. These options simply impact the format of the number as it is printed.

# Means Plot Tab

These options specify the three means plots.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Scaling

This option specifies whether the vertical axes of the three means plots are uniformly or separately scaled.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## Means Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

### Connect Line(s)

Click this box to connect the points for a particular factor. This makes it easier to spot patterns in the means.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the word 'Measurement.' Press the button on the right of the field to specify the font of the text.

# Data Plot Tab

These options specify the data plots.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Scaling

This option specifies whether the vertical axes of the three means plots are uniformly or separately scaled.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## Data Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the word 'Measurement.' Press the button on the right of the field to specify the font of the text.

# Residual Plot Tab

These options specify the residual plots.

## Vertical and Horizontal Axis

### Label

This is the text of the label. The characters *{Y}* and *{X}* are replaced by appropriate names. Press the button on the right of the field to specify the font of the text.

### Scaling

This option specifies whether the vertical axes of the three means plots are uniformly or separately scaled.

### Minimum

This option specifies the minimum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Maximum

This option specifies the maximum value displayed on the corresponding axis. If left blank, it is calculated from the data.

### Tick Label Settings…

Pressing these buttons brings up a window that sets the font, rotation, and number of decimal places displayed in the reference numbers along the vertical and horizontal axes.

### Major Ticks - Minor Ticks

These options set the number of major and minor tickmarks displayed on the axis.

### Show Grid Lines

This check box indicates whether the grid lines that originate from this axis should be displayed.

## Resid Plot Settings

### Plot Style File

Designate a scatter plot style file. This file sets all scatter plot options that are not set directly on this panel. Unless you choose otherwise, the default style file (Default) is used. These files are created in the Scatter Plot procedure.

## Titles

### Plot Title

This is the text of the title. The characters *{Y}* are replaced by the word 'Measurement' and the characters *{Z}* are replaced by the name of the group of data being plotted. Press the button on the right of the field to specify the font of the text.

# Symbols Tab

These options specify the attributes of the symbols used for each appraiser in the plots.

## Plotting Symbols

### Group 1 - 15

These options specify the symbols used in the plot of each appraiser. The first symbol is used by the first appraiser, the second symbol by the second appraiser, and so on. These symbols are provided to allow the various appraisers to be easily identified, even on black and white printers.

Clicking on a symbol box (or the small button to the right of the symbol box) will bring up a window that allows the color, width, and pattern of the line to be changed.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Running an R & R Study

This section presents an example of how to run an R & R study of the data that were displayed earlier in this chapter. These data are contained in the RRSTUDY database. In this example, ten chemical samples were selected for analysis. Each of three operators measured each of the ten samples three times. Each row contains one of the three trials for a particular sample. A trial consists of a measurement by each operator.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the R & R Study window.

**1 Open the RRSTUDY dataset.**

- From the **File** menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **RRSTUDY.S0**.
- Click **Open**.

**2 Open the R & R Study window.**

- On the menus, select **Analysis**, then **Quality Control**, then **R & R Study**. The R & R Study procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the R & R Study window, select the **Variables tab**.
- Double-click in the **Sample (Part) Variable** box. This will bring up the variable selection window.
- Select **Sample** from the list of variables and then click **Ok**.
- Double-click in the **Measurement Variable(s)** box. This will bring up the variable selection window.
- Select **Op1, Op2, Op3** from the list of variables and then click **Ok**.
- Enter **48** in the **Lower Spec Limit** box.
- Enter **88** in the **Upper Spec Limit** box. Note that 88-48 = 40 which is the tolerance.
- Enter **68** in the **Target Spec** box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Data Summary Section

**Data Summary Section**

| Item | Actual Count | Expected Count |
|---|---|---|
| Total Values | 90 | 90 |
| Sample | 10 | |
| Operators | 3 | |
| Replicates | 3 | |

This section presents a summary of the number of data values analyzed. In order for the analysis to be valid, the Actual Count must match the Expected Count in the Total Values row. When this

occurs, the design (data matrix) is said to be balanced. All of the formulas used are for balanced data matrices only.

### Total Values

The number of nonmissing data values in the dataset. If the design is balanced, the entry on this line equals the product of the entries on the next three lines.

### Samples

The number of samples (parts) found in the dataset.

### Appraisers

The number of appraiser (operator) variables selected.

### Replicates

The number of times an operator measured the same part.

## Expected Mean Square and Variance Component Section

**Expected Mean Square and Variance Component Section**

| Source Term | DF | Expected Mean Square | Variance Component | Lower 90% Conf. Limit | Upper 90% Conf. Limit |
|---|---|---|---|---|---|
| Sample (P) | 9 | R+3(PO)+9(P) | 5.615638 | 2.948817 | 15.33656 |
| Operators (O) | 2 | R+3(PO)+30(O) | 0.3563786 | 0.1016096 | 7.389713 |
| Interaction (PO) | 18 | R+3(PO) | 0.1251029 | 2.385001E-02 | 0.3455315 |
| Replicates (R) | 60 | R | 0.3444445 | 0.2613323 | 0.4785284 |

The expected mean square expressions and variance components are for each term in the analysis of variance model.

### Source Term

The source of variation or term in the model.

### DF

The degrees of freedom. The number of observations “used” by this term.

### Expected Mean Square

This is the symbolic value of the mean square for the term in the ANOVA model assuming balanced data (equal group counts). *“P”* represents $\sigma_P^2$. *“O”* represents $\sigma_O^2$. *“PO”* represents $\sigma_{PO}^2$. *“R”* represents $\sigma_E^2$.

### Variance Component

This is the expected value of corresponding variance in the ANOVA model assuming balanced data (equal group counts). Hence, the estimate of $\sigma_P^2$ is 5.615638 and the estimate of $\sigma_{PO}^2$ is 0.1251029. The formulas used for these estimates are

$$\sigma_O^2 = \frac{MS_O - MS_{PO}}{IK}$$

$$\sigma_P^2 = \frac{MS_P - MS_{PO}}{JK}$$

$$\sigma_{PO}^2 = \frac{MS_{PO} - MS_E}{K}$$

$$\sigma_E^2 = MS_E$$

where $MS_Q$ represents the mean square of term Q in an analysis of variance table.

### Lower (and Upper) Conf. Limit

These are the lower and upper confidence limits (interval estimate) of the variance components. The formulas used are found in Burdick and Larsen (1997). They are given as follows:

**Confidence Interval for $\sigma_O^2$ is**

$$Lower_O = \frac{MS_O - MS_{PO} - \sqrt{G_O^2 MS_O^2 + H_{PO}^2 MS_{PO}^2 + G_{O,PO} MS_O MS_{PO}}}{IK}$$

$$Upper_O = \frac{MS_O - MS_{PO} + \sqrt{H_O^2 MS_O^2 + G_{PO}^2 MS_{PO}^2 + H_{O,PO} MS_O MS_{PO}}}{IK}$$

**Confidence Interval for $\sigma_P^2$ is**

$$Lower_P = \frac{MS_P - MS_{PO} - \sqrt{G_P^2 MS_P^2 + H_{PO}^2 MS_{PO}^2 + G_{P,PO} MS_P MS_{PO}}}{JK}$$

$$Upper_P = \frac{MS_P - MS_{PO} + \sqrt{H_P^2 MS_P^2 + G_{PO}^2 MS_{PO}^2 + H_{S,PO} MS_P MS_{PO}}}{JK}$$

**Confidence Interval for $\sigma_{PO}^2$ is**

$$Lower_{PO} = \frac{MS_{PO} - MS_E - \sqrt{G_{PO}^2 MS_{PO}^2 + H_E^2 MS_E^2 + G_{PO,E} MS_{PO} MS_E}}{K}$$

$$Upper_{PO} = \frac{MS_{PO} - MS_E + \sqrt{H_{PO}^2 MS_{PO}^2 + G_E^2 MS_E^2 + H_{PO,E} MS_{PO} MS_E}}{K}$$

**Confidence Interval for $\sigma_E^2$ is**

$$Lower_E = (1 - G_E) MS_E$$

$$Upper_E = (1 + H_E) MS_E$$

where

$$G_q = 1 - \frac{1}{F_{\alpha, n_q, \infty}}$$

$$H_q = \frac{1}{F_{1-\alpha,n_q,\infty}} - 1$$

$$G_{qr} = \frac{\left(F_{\alpha,n_q,n_r} - 1\right)^2 - G_q^2 F_{\alpha,n_q,n_r}^2 - H_r^q}{F_{\alpha,n_q,n_r}}$$

$$H_{qr} = \frac{\left(1 - F_{1-\alpha,n_q,n_r}\right)^2 - H_q^2 F_{1-\alpha,n_q,n_r}^2 - G_r^q}{F_{1-\alpha,n_q,n_r}}$$

and $F_{\alpha,n_q,n_r}$ is the F distribution with an area equal to $\alpha$ to the right. The confidence level of these intervals is $100(1-2\alpha)\%$ . The subscripts *q* and *r* refer to the terms *O*, *P*, *PO*, and *E*. The *n's* are given by

$$n_S = I - 1$$

$$n_O = J - 1$$

$$n_{SO} = (I-1)(J-1)$$

$$n_E = IJ(K-1)$$

## Analysis of Variance Section

**Analysis of Variance Section**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level |
|---|---|---|---|---|---|
| Sample | 9 | 461.3445 | 51.26049 | 71.22 | 0.000000 |
| Operators | 2 | 22.82222 | 11.41111 | 15.85 | 0.000107 |
| Interaction | 18 | 12.95556 | 0.7197531 | 2.09 | 0.017450 |
| Replicates | 60 | 20.66667 | 0.3444445 | | |
| Total (Adjusted) | 89 | 517.7889 | | | |
| Total | 90 | | | | |

### Source Term

The source of variation. The term in the model.

### DF

The degrees of freedom. The number of observations "used" by the corresponding model term.

### Sum of Squares

This is the sum of squares for this term. It is usually included in the ANOVA table for completeness, not for direct interpretation.

### Mean Square

An estimate of the variation accounted for by this term. The sum of squares divided by the degrees of freedom.

### F-Ratio

The ratio of the mean square for this term and the mean square of its corresponding error term. This is also called the F-test value.

### Prob Level

The significance level of the above F-ratio. The probability of an F-ratio larger than that obtained by this analysis. For example, to test at an alpha level of 0.05, this probability would have to be less than 0.05 to make the F-ratio significant. Note that if the value is significant at the specified value of alpha, a star is placed to the right of the F-Ratio.

## Variance Section

**Variance Section**

| Term | Variance | % Total Variance | Standard Deviation | Lower 90% Conf. Limit of Std Dev | Upper 90% Conf. Limit of Std Dev | % Total Std Dev |
|---|---|---|---|---|---|---|
| Sample | 5.615638 | 87.1782 | 2.3697 | 1.7172 | 3.9162 | 93.3693 |
| Operators | 0.356379 | 5.5325 | 0.5970 | 0.3188 | 2.7184 | 23.5212 |
| Interaction | 0.125103 | 1.9421 | 0.3537 | 0.1544 | 0.5878 | 13.9360 |
| Reproducibility | 0.481481 | 7.4746 | 0.6939 | 0.4349 | 2.7415 | 27.3397 |
| Repeatability | 0.344444 | 5.3472 | 0.5869 | 0.5112 | 0.6918 | 23.1241 |
| R and R | 0.825926 | 12.8218 | 0.9088 | 0.7443 | 2.8044 | 35.8076 |
| Total Variation | 6.441564 | 100.0000 | 2.5380 | 1.9394 | 4.2947 | 100.0000 |

This report presents estimates of the variance and standard deviation of various terms of interest in an R & R study.

### Term

These are the names of the variance terms being estimated. The first few terms were discussed above in the Expected Mean Square and Variance Component Report. “Sample” refers to $\sigma_P^2$, the variability between samples (parts). “Operators” refers to $\sigma_O^2$, the variability between appraisers (operators). “Interaction” refers to $\sigma_{PO}^2$, the interaction variation. “Repeatability” refers to $\sigma_E^2$, the variability that occurs when one appraiser measures the same sample over and over.

“Reproducibility” refers to the variation among appraisers which is $\gamma_1 = \sigma_O^2 + \sigma_{PO}^2$. “R and R” refers to the sum of Reproducibility and Repeatability which is $\gamma_2 = \sigma_E^2 + \sigma_O^2 + \sigma_{PO}^2$. “Total Variation” is the sum of all four sources of variation $\sigma_T^2 = \sigma_E^2 + \sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2$.

### Variance

These are the estimated values of the variances of the terms listed above. The formulas for the first four terms were given in the Expected Mean Square and Variance Component Report. The formulas for the last three items are as follows.

$$\hat{\gamma}_1 = \frac{MS_O + (I-1)MS_{PO} - I\left(MS_E\right)}{IK}$$

$$\hat{\gamma}_2 = \frac{MS_O + (I-1)MS_{PO} + I\left(K-1\right)MS_E}{IK}$$

$$\hat{\sigma}_T^2 = \frac{MS_P}{(I-1)(J-1)^2} + \frac{MS_O}{(I-1)^2(J-1)} + \left[\frac{(I-1)(J-1)-(I-1)-(J-1)}{(I-1)^2(J-1)^2}\right] MS_{PO}$$
$$+ \left[\frac{(I-1)(J-1)-1}{(I-1)(J-1)}\right] MS_E$$

### % Total Variance

This shows the percentage that each variance is of the Total Variation variance.

### Standard Deviation

This is the square root of the variance.

### Lower (and Upper) 90% Conf. Limit

These are the lower and upper confidence limits (interval estimate) for the standard deviation shown in the previous column. The formulas used are found by taking the square root of the corresponding variance confidence limits found in Burdick and Larsen (1997). The values of the first four terms were given in the Expected Mean Square and Variance Component Report. The formulas for the last three items are as follows.

**Confidence Interval for Reproducibility, $\sigma_O^2 + \sigma_{PO}^2$, is**

$$Lower_{\gamma_1} = \hat{\gamma}_1 - \sqrt{V_{L\gamma}}$$

$$Upper_{\gamma_1} = \hat{\gamma}_1 + \sqrt{V_{U\gamma}}$$

where

$$V_{L\gamma} = \frac{\begin{array}{l} G_O^2 MS_O^2 + G_{PO}^2 (I-1)^2 MS_{PO}^2 + H_E^2 I^2 MS_E^2 + G_{O,E}(I) MS_O MS_E \\ + G_{PO,E} I(I-1) MS_{PO} MS_E + G_{O,PO}^*(I-1) MS_O MS_{PO} \end{array}}{(IK)^2}$$

$$V_{U\gamma} = \frac{\begin{array}{l} H_O^2 MS_O^2 + H_{PO}^2 (I-1)^2 MS_{PO}^2 + G_E^2 I^2 MS_E^2 + H_{O,E}(I) MS_O MS_E \\ + H_{PO,E} I(I-1) MS_{PO} MS_E \end{array}}{(IK)^2}$$

and

$$G_{O,PO}^* = \left(1 - \frac{1}{F_{\alpha, I(J-1), \infty}}\right)^2 \frac{I^2}{I-1} - \frac{G_O^2}{I-1} - (I-1) G_{PO}^2$$

**Confidence Interval for R and R, $\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2$, is**

$$Lower_{\gamma_2} = \hat{\gamma}_2 - \sqrt{V_{LRR}}$$

$$Upper_{\gamma_2} = \hat{\gamma}_2 + \sqrt{V_{URR}}$$

where

$$V_{LRR} = \frac{G_O^2 MS_O^2 + G_{PO}^2 (I-1)^2 MS_{PO}^2 + G_E^2 I^2 (K-1)^2 MS_E^2}{(IK)^2}$$

$$V_{URR} = \frac{H_O^2 MS_O^2 + H_{PO}^2 (I-1)^2 MS_{PO}^2 + G_E^2 I^2 (K-1)^2 MS_E^2}{(IK)^2}$$

**Confidence Interval for Total Variation, $\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2$, is**

$$Lower_T = \hat{\sigma}_T^2 - \sqrt{V_{LT}}$$

$$Upper_T = \hat{\sigma}_T^2 + \sqrt{V_{UT}}$$

where

$$V_{LT} = G_P^2 C_P^2 MS_P^2 + G_O^2 C_O^2 MS_O^2 + G_{PO}^2 C_{PO}^2 MS_{PO}^2 + G_E^2 C_E^2 MS_E^2$$

$$V_{UT} = H_P^2 C_P^2 MS_P^2 + H_O^2 C_O^2 MS_O^2 + H_{PO}^2 C_{PO}^2 MS_{PO}^2 + H_E^2 C_E^2 MS_E^2$$

and

$$C_P = \frac{1}{(I-1)(J-1)^2}$$

$$C_O = \frac{1}{(I-1)^2(J-1)}$$

$$C_{PO} = \frac{(I-1)(J-1)-(I-1)-(J-1)}{(I-1)^2(J-1)^2}$$

$$C_E = \frac{(I-1)(J-1)-1}{(I-1)(J-1)}$$

### % Total Std Dev

This column gives the percentage that each standard deviation is of the total standard deviation. Because the total standard deviation is not equal to the sum of the individual standard deviations (it is the variances that are summed), these percentages may total to more than 100.

# Percent of Process Variation R & R Section

**Percent of Process Variation R & R Section**

| Term | Lower 90% Conf. Limit | 5.15 Std Dev | Upper 90% Conf. Limit | % Total Variation | Percent Contribution |
|---|---|---|---|---|---|
| Sample | 8.8436 | 12.2041 | 20.1684 | 93.3693 | 87.1782 |
| Operator | 1.6416 | 3.0744 | 13.9998 | 23.5212 | 5.5325 |
| Interaction | 0.7953 | 1.8215 | 3.0273 | 13.9360 | 1.9421 |
| Reproducibility | 2.2395 | 3.5735 | 14.1187 | 27.3397 | 7.4746 |
| Repeatability | 2.6327 | 3.0225 | 3.5626 | 23.1241 | 5.3472 |
| R and R | 3.8332 | 4.6803 | 14.4425 | 35.8076 | 12.8218 |
| Total Variation | 9.9877 | 13.0708 | 22.1175 | 100.0000 | 100.0000 |

Since the % R & R value is greater than 30%, the measurement system is not acceptable.
Identify the measurement problems and correct them.

This report gives components of the process variation scaled by multiplying by the Sigma Multiplier value (which defaults to 5.15). This multiplication puts all values in same metric as the specification limits so they can be compared directly. For example, the variability that occurs when the same appraiser measures the same sample twice adds between 2.6327 and 3.5626 to the measurement standard deviation. Hence, by comparing these values, we can see the relative impact of each source of variation.

### Term

These are the names of the terms being displayed. All of these terms have been defined previously.

### Lower (and Upper) 90% Conf. Limit

These are the lower and upper confidence limits (interval estimate) for the standard deviation shown in between these two columns. The formulas used are found by taking the square root of the corresponding variance confidence limits found in Burdick and Larsen (1997). The values are multiplied by the Sigma Multiplier as discussed above.

### 5.15 Std Dev

This is the square root of the variance associated with each term multiplied by the Sigma Multiplier (5.15 is the default).

### % Total Variation

This is 100 times the ratio of this term’s standard deviation to the total variation’s standard deviation. One of the key statistics to look at is whether the R and R value in this column is small enough. If the R and R value is less than 10%, the measurement procedure is deemed excellent. When it is less than 20%, it is deemed adequate. When it is less than 30%, it is marginal. When the R and R value is greater than 30%, it should not be used for process monitoring.

### Percent Contribution

This is 100 times the ratio of this term’s variance to the total variation’s variance.

# Percent of Tolerance R & R Section

**Percent of Tolerance R & R Section**

| Term | Lower 90% Conf. Limit | 5.15 Std Dev | Upper 90% Conf. Limit | Percent Tolerance |
|---|---|---|---|---|
| Sample | 8.8436 | 12.2041 | 20.1684 | 30.5103 |
| Operator | 1.6416 | 3.0744 | 13.9998 | 7.6860 |
| Interaction | 0.7953 | 1.8215 | 3.0273 | 4.5539 |
| Reproducibility | 2.2395 | 3.5735 | 14.1187 | 8.9338 |
| Repeatability | 2.6327 | 3.0225 | 3.5626 | 7.5563 |
| R and R | 3.8332 | 4.6803 | 14.4425 | 11.7009 |
| Total Variation | 9.9877 | 13.0708 | 22.1175 | 32.6771 |

| | |
|---|---|
| Upper Spec Limit | 88 |
| Lower Spec Limit | 48 |
| Tolerance | 40 |

Since the % R & R value is between 10% and 20%, the measurement system is acceptable.

This report is similar to the last report, except that the denominator of the percentages in the last column is the tolerance rather than the total variation.

### Term

These are the names of the terms being displayed. All of these terms have been defined previously.

### Lower (and Upper) 90% Conf. Limit

These are the lower and upper confidence limits (interval estimate) for the standard deviation shown in between these two columns. The formulas used are found by taking the square root of the corresponding variance confidence limits found in Burdick and Larsen (1997). The values are multiplied by the Sigma Multiplier as discussed above.

### 5.15 Std Dev

This is the square root of the variance associated with each term multiplied by the Sigma Multiplier (5.15 is the default).

### Percent Tolerance

This is 100 times 5.15 times the ratio of this term’s standard deviation to the tolerance. One of the key statistics to look at is whether the R and R value in this column is small enough. If the R and R value is less than 10%, the measurement procedure is deemed excellent. When it is less than 20%, it is deemed adequate. When it is less than 30%, it is marginal. When the R and R value is greater than 30%, it should not be used for process monitoring.

### Upper (Lower) Spec Limits and Tolerance

The upper and lower specification limits are specified by the user. The tolerance is the upper specification limit minus the lower specification limit.

# R & R Indices Section

**R & R Indices Section**

| Index | Lower 90% Conf. Limit | Value | Upper 90% Conf. Limit |
|---|---|---|---|
| Distinct Categories | 1.1924 | 3.6876 | 6.2979 |
| Signal-to-Noise Ratio | 0.8431 | 2.6075 | 4.4533 |
| Measurement Error | 5.5823 | 6.8160 | 21.0328 |
| Precision-to-Tolerance | 11.1647 | 13.6321 | 42.0655 |

Since the lower confidence limit of Distinct Categories is less than 3,
the measurement process may be inadequate.
Since the upper confidence limit of Measurement Error is less than 25%,
measurement error can be ignored in decision making.

This report gives values with confidence limits for four indices that have been found useful in analyzing R & R data. You will have to decide whether to use the point estimate (the Value) or the interval estimate (the Confidence Limits) when making decisions.

The first three statistics on this report are based on the ratio

$$\delta = \frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}$$

Confidence limits for this ratio are given below.

### Single-to-Noise Ratio

This index is given by the formula

$$SNR = \sqrt{\hat{\delta}} = \sqrt{\frac{\hat{\sigma}_P^2}{\hat{\sigma}_O^2 + \hat{\sigma}_{PO}^2 + \hat{\sigma}_E^2}}$$

As you can see, it is the ratio of the sample-to-sample standard deviation and the measurement (R and R) variation. As a manufacturer, we are really interested in the sample-to-sample variability. The measurement standard deviation estimates the noise that is added to the sample-to-sample variability by the approximate nature of the measurement system.

The *Measurement Systems Analysis Reference Manual* (AGIG 1995) recommend that this value exceed 2.12.

### Distinct Categories

This index is the number of distinct product categories that can be reliably distinguished by the measurement procedure. Its formula is

$$Distinct\ Categories = \sqrt{2\hat{\delta}} = \sqrt{\frac{2\hat{\sigma}_P^2}{\hat{\sigma}_O^2 + \hat{\sigma}_{PO}^2 + \hat{\sigma}_E^2}}$$

The *Measurement Systems Analysis Reference Manual* (AGIG 1995) recommend that this value exceed 3.

### Measurement Error

This index compares the measurement standard deviation to the tolerance, where tolerance is the difference between the upper specification limit (USL) and lower specification limit (LSL). The value is calculated using the formula

$$M = \frac{3\sqrt{\hat{\sigma}_O^2 + \hat{\sigma}_{PO}^2 + \hat{\sigma}_E^2}}{USL - LSL} \times 100\%$$

A rule-of-thumb is that this value should be less than 25% in order for the measurement system to be deemed adequate.

### Precision-to-Tolerance

A slightly different version of the Measurement Error index is the Precision-to-Tolerance ratio (P/T) which is defined as

$$PT = \frac{6\sqrt{\hat{\sigma}_O^2 + \hat{\sigma}_{PO}^2 + \hat{\sigma}_E^2}}{USL - LSL} \times 100\%$$

### Confidence Limits for Ratio

The first three statistics on this report are function of the ratio

$$\delta = \frac{\sigma_P^2}{\sigma_O^2 + \sigma_{PO}^2 + \sigma_E^2}$$

The formulae for confidence limits of this statistic are given by Burdick and Larsen (1997). They are included here for easy reference. The approximate $100(1-2\alpha)\%$ confidence limits are

$$\left[L_\delta; U_\delta\right]$$

where

$$U_\delta = \frac{I}{J}\left[\frac{B_U + \sqrt{Q_U}}{2A_U}\right]$$

$$L_\delta = \frac{I}{J}\left[\frac{B_L - \sqrt{Q_L}}{2A_L}\right]$$

$$Q_U = Max\left[0, B_U^2 - 4A_U C_U\right]$$

$$Q_L = Max\left[0, B_L^2 - 4A_L C_L\right]$$

$$\begin{aligned} A_U = {} & \left(1-G_O^2\right)MS_O^2 + \left(I-1\right)^2\left(1-G_{PO}^2\right)MS_{PO}^2 + I^2\left(K-1\right)^2\left(1-G_E^2\right)MS_E^2 \\ & + 2\left(I-1\right)MS_O MS_{PO} + 2I\left(K-1\right)MS_O MS_E + 2I\left(I-1\right)\left(K-1\right)MS_{PO} MS_E \end{aligned}$$

$$\begin{aligned} A_L = {} & \left(1-H_O^2\right)MS_O^2 + \left(I-1\right)^2\left(1-H_{PO}^2\right)MS_{PO}^2 + I^2\left(K-1\right)^2\left(1-H_E^2\right)MS_E^2 \\ & + 2\left(I-1\right)MS_O MS_{PO} + 2I\left(K-1\right)MS_O MS_E + 2I\left(I-1\right)\left(K-1\right)MS_{PO} MS_E \end{aligned}$$

$$\begin{aligned} B_U = {} & -2\left(I-1\right)\left(1-G_{PO}^2\right)MS_{PO}^2 + \left(2+H_{P,O}\right)MS_P MS_O + \left(I-1\right)\left(2+H_{P,PO}\right)MS_P MS_{PO} \\ & + I\left(K-1\right)\left(2+H_{P,E}\right)MS_P MS_E - 2MS_O MS_{PO} - 2I\left(K-1\right)MS_{PO} MS_E \end{aligned}$$

$$B_L = -2(I-1)\left(1-H_{PO}^2\right)MS_{PO}^2 + \left(2+G_{P,O}\right)MS_P MS_O + (I-1)\left(2+G_{P,PO}\right)MS_P MS_{PO}$$
$$+ I(K-1)\left(2+G_{P,E}\right)MS_P MS_E - 2MS_O^2 MS_{PO} - 2I(K-1)MS_{PO} MS_E$$

$$C_U = \left(1-H_P^2\right)MS_P^2 + \left(1-G_{PO}^2\right)MS_{PO}^2 - \left(2+H_{P,PO}\right)MS_P MS_{PO}$$

$$C_L = \left(1-G_P^2\right)MS_P^2 + \left(1-H_{PO}^2\right)MS_{PO}^2 - \left(2+G_{P,PO}\right)MS_P MS_{PO}$$

## Means and Bias Section

**Means and Bias Section**

| Term | Count | Mean | Deviation From Target |
|---|---|---|---|
| Overall | 90 | 68.589 | 0.589 |
| Sample | | | |
| 1 | 9 | 67.667 | -0.333 |
| 2 | 9 | 66.667 | -1.333 |
| 3 | 9 | 68.333 | 0.333 |
| 4 | 9 | 68.000 | 0.000 |
| 5 | 9 | 68.889 | 0.889 |
| 6 | 9 | 69.667 | 1.667 |
| 7 | 9 | 67.778 | -0.222 |
| 8 | 9 | 74.778 | 6.778 |
| 9 | 9 | 67.889 | -0.111 |
| 10 | 9 | 66.222 | -1.778 |
| Operators | | | |
| Op1 | 30 | 67.967 | -0.033 |
| Op2 | 30 | 69.200 | 1.200 |
| Op3 | 30 | 68.600 | 0.600 |
| Sample, Operators | | | |
| 1,Op1 | 3 | 67.667 | -0.333 |
| 1,Op2 | 3 | 67.333 | -0.667 |
| 1,Op3 | 3 | 68.000 | 0.000 |
| 2,Op1 | 3 | 66.333 | -1.667 |
| 2,Op2 | 3 | 67.333 | -0.667 |
| 2,Op3 | 3 | 66.333 | -1.667 |
| 3,Op1 | 3 | 67.667 | -0.333 |
| 3,Op2 | 3 | 69.333 | 1.333 |
| 3,Op3 | 3 | 68.000 | 0.000 |
| 4,Op1 | 3 | 67.000 | -1.000 |
| 4,Op2 | 3 | 69.333 | 1.333 |
| 4,Op3 | 3 | 67.667 | -0.333 |
| 5,Op1 | 3 | 68.000 | 0.000 |
| 5,Op2 | 3 | 70.000 | 2.000 |
| 5,Op3 | 3 | 68.667 | 0.667 |
| 6,Op1 | 3 | 68.667 | 0.667 |
| 6,Op2 | 3 | 70.333 | 2.333 |
| 6,Op3 | 3 | 70.000 | 2.000 |
| 7,Op1 | 3 | 67.000 | -1.000 |
| 7,Op2 | 3 | 68.000 | 0.000 |
| 7,Op3 | 3 | 68.333 | 0.333 |
| 8,Op1 | 3 | 74.333 | 6.333 |
| 8,Op2 | 3 | 75.000 | 7.000 |
| 8,Op3 | 3 | 75.000 | 7.000 |
| 9,Op1 | 3 | 67.000 | -1.000 |
| 9,Op2 | 3 | 68.667 | 0.667 |
| 9,Op3 | 3 | 68.000 | 0.000 |
| 10,Op1 | 3 | 66.000 | -2.000 |
| 10,Op2 | 3 | 66.667 | -1.333 |
| 10,Op3 | 3 | 66.000 | -2.000 |

The main purpose of this report is to acquaint you with the data and allow you to quickly find outliers. We will discuss more about outliers below.

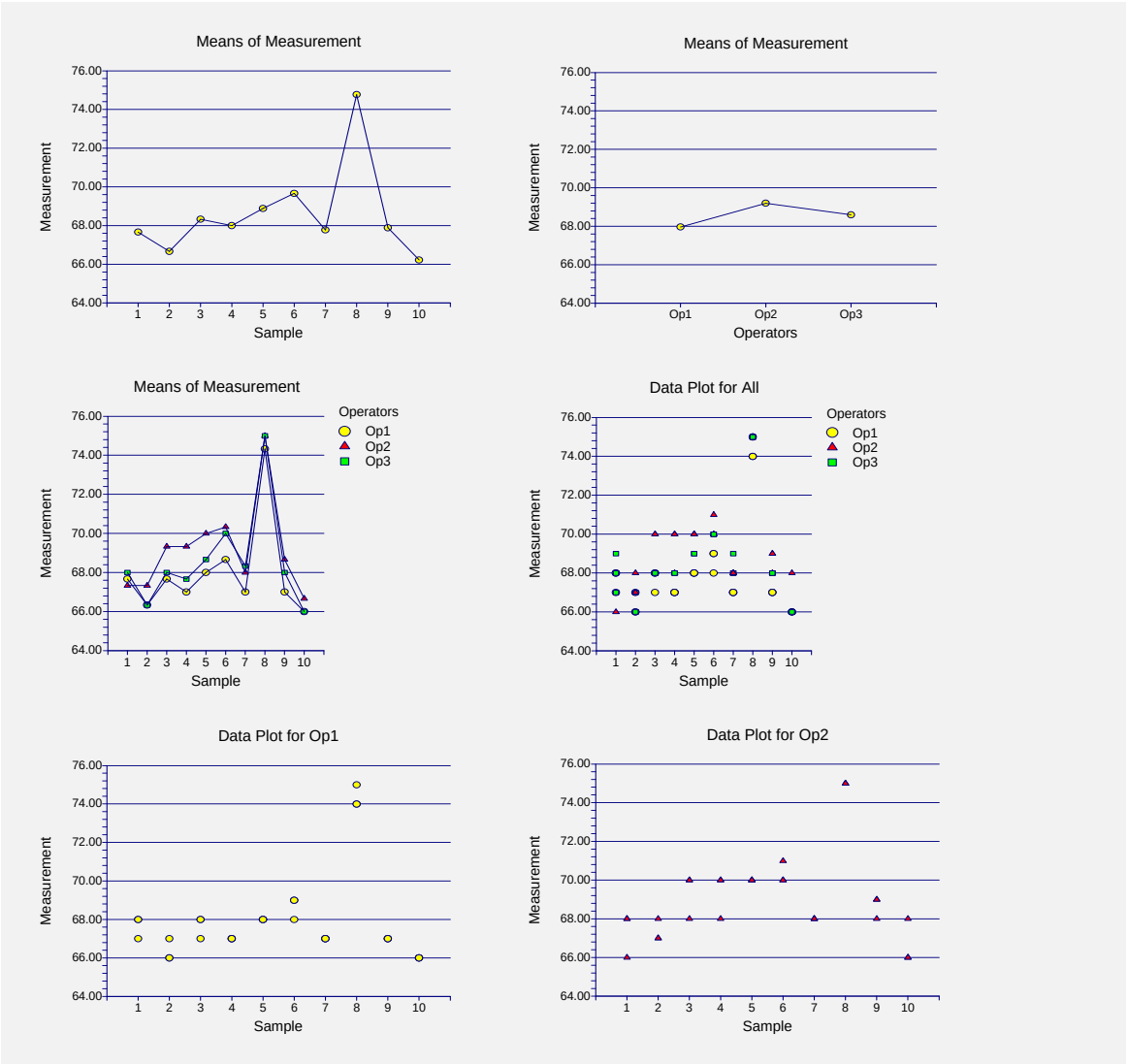
### Term
The label for this line of the report.

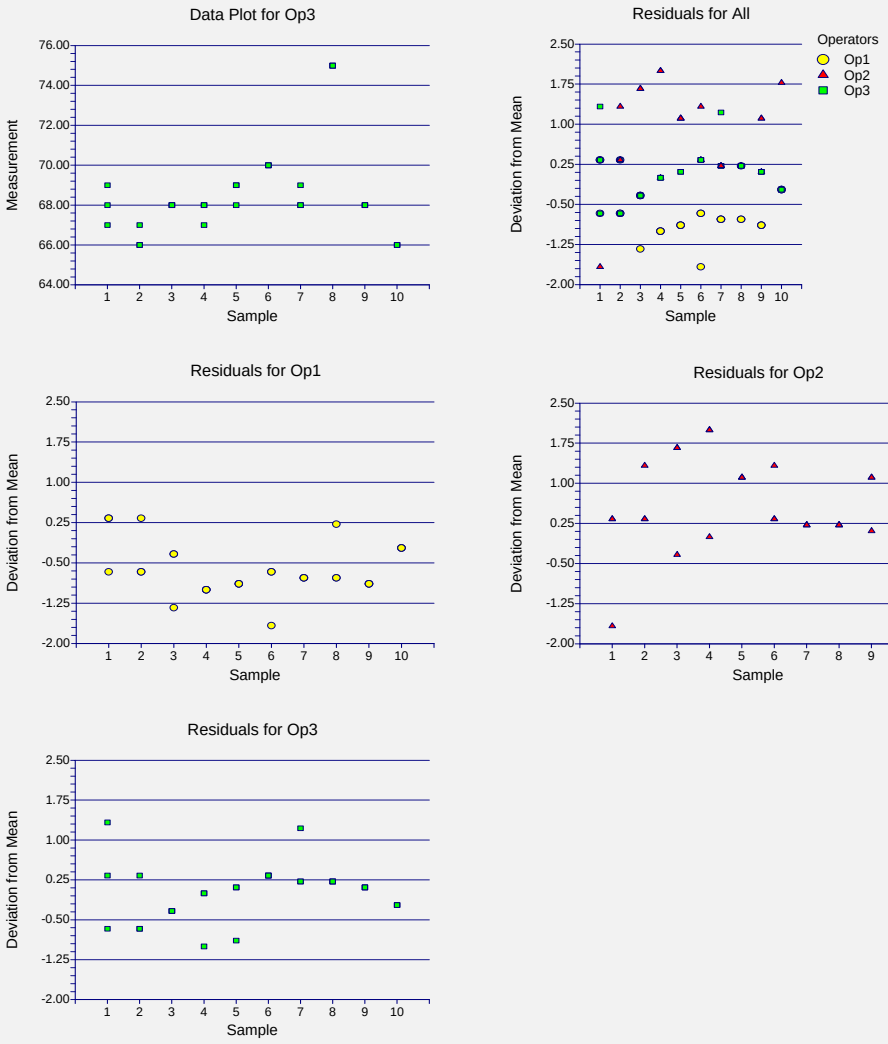
### Count
The number of observations in the mean.

### Mean
The value of the sample mean.

### Bias
This is the difference between the mean and the Target Spec.

## Plots Section






This section displays various plots of means, the original data, and residuals. You should look through these plots for unexpected patterns, trends, and outliers.

The plots of the means let you analyze the systematic variation in your data. For example, you can see whether one appraiser was very different from the rest. You can also determine whether certain samples were extremely different from the others.

The data plots let you see the original data. In these plots, you will be able to quickly find outliers (which often turn out to be data entry errors) and unusual patterns. This plot will give you a good feel for the variation in your data.

The residual plots show the deviation between each data value and the sample (part) mean for that value. This lets you view the measurement error.

## Example 2 – Analysis of Variance Data

In this example, the RRSTUDY database has been reformatted to match the more typical data format necessary to run an analysis of variance on the data. The difference is that the operator factor is explicitly represented as a variable and only one measurement is given per row. This

format requires ninety rows instead of thirty. The first six rows are displayed here. The complete dataset is contained in the RRSTUDY1 database.

**RRSTUDY1 dataset (subset)**

| Sample | Operator | Measurement |
|---|---|---|
| 1 | Op1 | 67 |
| 1 | Op1 | 68 |
| 1 | Op1 | 68 |
| 2 | Op1 | 67 |
| 2 | Op1 | 66 |
| 2 | Op1 | 66 |

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the R & R Study window.

**1 Open the RRSTUDY1 dataset.**

- From the **File** menu of the NCSS Data window, select **Open**.
- Select the **Data** subdirectory of your NCSS directory.
- Click on the file **RRSTUDY1.S0**.
- Click **Open**.

**2 Open the R & R Study window.**

- On the menus, select **Analysis**, then **Quality Control**, then **R & R Study**. The R & R Study procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the variables.**

- On the R & R Study window, select the **Variables tab**.
- Double-click in the **Sample (Part) Variable** box. This will bring up the variable selection window.
- Select **Sample** from the list of variables and then click **Ok**.
- Double-click in the **Appraiser (Operator) Variable** box. This will bring up the variable selection window.
- Select **Operator** from the list of variables and then click **Ok**.
- Double-click in the **Measurement Variable(s)** box. This will bring up the variable selection window.
- Select **Measurement** from the list of variables and then click **Ok**.
- Enter **48** in the **Lower Spec Limit** box.
- Enter **88** in the **Upper Spec Limit** box. Note that 88-48 = 40 which is the tolerance.
- Enter **68** in the **Target Spec** box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

Since the same data are being analyzed, the reports are the same as in Example 1.

# Chapter 260

# Two-Level Designs

## Introduction

This program generates a $2^k$ factorial design for up to seven factors. It allows the design to be blocked and replicated. The design rows may be output in standard or random order. The output of this program will be to the current database with the data from the specified design. Hence, this particular program does not analyze data, it generates it.

When blocking is specified, the program checks to see if the design is listed on page 408 of Box and Hunter (1978). If it is one of the designs specified there, the indicated confounding pattern is used. If not, the blocks are confounded using the standard procedure in which highest-order interactions are confounded first, so long as they do not cause main effects to be confounded with blocks. The blocking pattern is reported by the analysis program, so it is not reported by this program.

## Experimental Design

Experimental design is the planning of an efficient, reliable, and accurate technical study. The range of application of experimental design principles is as broad as science and industry. One person may be planning a long-term agricultural experiment, while another may have eight hours to rectify a production problem. How can we expect that the same methods are appropriate in all situations?

Of course, we cannot. Through the years, researchers and statisticians working together have outlined the basic steps necessary to conduct an effective investigation. These steps form an experimental strategy that seems to work well in many settings.

The experimental design modules lend you, the investigator, a hand with the planning and analysis of your investigation. Once you have determined the scope of your investigation, the design modules will provide a data collection plan that will minimize the amount of data collected and maximize the amount of conclusive information available. They will also provide a statistical analysis of your experimental results after the data have been collected.

The experimental design chapters will not attempt to teach you the principles of experimental design. There are many excellent books and pamphlets on this subject. The focus of the manual will be to remind you of the basic principles of experimental design and then explain where and how the program can help in your study. We suggest that you consult one or two of the following texts for detailed coverage of experimental design: Box, Hunter, and Hunter (1978), Davies (1956), Lawson (1987), or Montgomery (1984).

# The Role of Statistics in Science

Statistics has been called the science of science. The scientific method consists of developing a theory or hypothesis, determining the consequences of this theory, and then comparing these consequences with facts (already available or determined from experimentation). When facts are found that contradict the theory, the theory must be modified, the consequences again determined, and all facts reconsidered.

The field of statistics is used in two phases of the scientific method. First, statistical design principles are used in the planning phase to determine an efficient and accurate method for collecting data (facts). Second, statistical analysis techniques are used to determine if the data are compatible with the proposed theory. Tools are provided for both of these phases in our statistical package.

# Experimental Design Definitions

## Alias

Two terms are aliased if their levels are identical throughout the design (except possibly for a difference in sign). Aliasing occurs in designs that are less than one full replication. The two terms are completely confounded with one another. It is impossible to determine from the data if an effect is due to the first, second, or both terms.

## Blocking

A block refers to a batch of runs conducted together. For example, a block may be the experiments run on a particular day, or the experiments conducted on a particular batch of material.

## Confounding

Two terms are confounded when their influences on the response variable cannot be separated. Confounding usually occurs when blocks are equated to high-order interactions.

## Experiment (Run)

An action to at least one of the items being studied which has an observable outcome. Each run produces one observation (value) of the response variable.

## Experimental Design

The collection of experiments to be completed during an investigation or study.

## Experimental Error

The influence on the response of all independent variables not included in the study. This *error* is a fact of life, since it is usually impossible to control every independent variable that might influence the response.

## Factorial Designs

A factorial design consists of all combinations of factor levels of two or more factors. The designs we generate all have factors with two, three, or five levels. Most of the designs are two-level designs. Since the total number of factor-level combinations is the product of the number of levels of each factor, these two-level designs are known as $2^k$ factorial designs (where k is the number of factors).

The two levels of each factor are often referred to as the high and the low levels. For example, if one of the factors were agitation at 100 rpm and 200 rpm, then 100 would be the low level and 200 would be the high level.

The designs produced by this program are orthogonal. This means that an equal amount of information is provided about the influence of each factor. It also means that there is no overlapping of information. The study clearly shows the unique influence of each factor.

One of the greatest strengths of the factorial experiment is that it allows the study of several factors at once, rather than only one factor at a time. Since each factor is paired with all possible combinations of the other factors, the researcher is confident that the measured effect of the factor is valid under a broad range of conditions.

## Independent Variable (Factor)

A variable whose influence on the response variable is being studied by deliberately varying it from run to run.

## Interaction

The interaction among factors refers to that part of the change in the response from run to run that may be accounted for by a specific combination of two or more factors. Another way of explaining interaction is that the average effect of one factor depends on the level of another factor.

The order of an interaction is the number of factors in the interaction. Hence AB is a second-order interaction and ABCD is a fourth-order interaction.

The Taylor's series expansion of a function is often used to justify the assumption that higher-order interactions are less significant (smaller influence on the response) than are main effects and low-order interactions.

## Levels

A factor (independent variable) is set at different values or levels during an experiment.

## Main Effect

The change in the average response as a factor is varied is called the main effect of that factor. In a factor with two levels, the main effect is the average of all runs at the high level of the factor minus the average of all runs at the low level of the factor.

## Response or Dependent Variable

The variable whose value is observed at the completion of each run.

## Replication

This is the number of times an experiment is repeated at identical factor levels. You must have some replication to determine the underlying (error) variability that occurs in the experiment. One *rep* refers to the running of every possible factor combination. Designs may be partially replicated (a few treatment settings are repeated), fractionally replicated (less than one complete replication), or completely replicated. It should be obvious that each time a run is repeated, the precision of the experimental results is increased.

## Two-Level Factorial Designs

All of the designs provided are factorial designs. Two-level designs are those in which all factors have only two values. This may seem like a severe restriction, but in many studies, this is all that is needed.

Factorial designs allow you to fit linear (as opposed to quadratic) models with all possible interactions. The number of runs is often quite large, so the runs are often grouped together in blocks.

## Fractional Factorial Designs

Fractional factorial designs are constructed by taking well-chosen subsets of a complete factorial design. Fractional factorials are useful because they require much fewer runs, although they do not allow the separation of main effects from high-order interactions.

This program gives two-level fractional factorial designs. These are usually defined as one-half rep, one-quarter rep, etc. They may be run all at once or in blocks.

## Screening Designs

Screening designs are used in the initial phases of a study when you wish to investigate the main effects of several factors (up to 31) simultaneously. These designs allow you to determine which factors warrant closer investigation and which may be ignored.

Screening designs allow the investigation of main effects only. They use a small fraction of the total runs that would be needed for a complete factorial design.

Many of the Taguchi designs are really screening designs.

## Response Surface Designs

The program provides Central Composite and Box-Behnken response surface designs. These designs provide for factors with more than two levels.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Simulated Response Variable

This optional variable will contain a computer-simulated response value for each row. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

#### Block Variable

The variable to contain the block identification numbers. The blocks are numbered from one to B, where B is the number of blocks. This variable is optional. If this option is left blank, no blocks will be generated.

#### First Factor Variable

This is where the group of variables that is to contain your design begins. The K-1 variables after this variable are also filled with data. The number of variables used is determined by the number of Factor Values boxes that contain data.

Warning: The program fills these variables with data, so any previous data will be lost.

### Data Storage Variables – Storage Options

#### Sort Order

The order of the generated rows. The rows may be in random or standard order.

- **Random**

  The rows are randomly ordered (random blocks and random rows within blocks). Use this option when the order of application to experimental units is governed by the row number.

- **Standard**

  The rows are not randomly ordered. Instead, they are placed in standard order. Use this option when you want to quickly see the structure of the design.

## Experimental Setup

### Replications

The number of replications (repeats) of the entire experiment.

### Block Size

The number of experiments (runs) per block. This determines the number of blocks. This number must be a power of 2 (2, 4, 8, 16, etc.)

### Factor Values

Each factor has two possible values (levels) which are specified here. These are the values that will be written to the database. The first value is used to represent the low value. The second value represents the high value. You may use both text and numeric values.

The number of variables created depends on how many of these boxes have values in them.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Two-Level Design

This section presents an example of how to generate an experimental design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

In this example, we will show you how to generate a five-factor design in blocks of eight runs each. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Two-Level Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Two-Level Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Two-Level Designs**. The Two-Level Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Two-Level Designs window, select the **Design tab**.
- Enter **1** in the **Simulated Response Variable** box.
- Enter **2** in the **Block Variable** box.
- Enter **3** in the **First Factor Variable** box.
- Select **Standard** in the **Sort Order** box.
- Select **1** in the **Replications** box.
- Select **8** in the **Block Size** box.
- Set the **first Factor Value** box to **1 2**.
- Set the **second Factor Value** box to **10 20**.
- Set the **third Factor Value** box to **Low High**.
- Set the **fourth Factor Value** box to **-1 1**.
- Set the **fifth Factor Value** box to **0 1**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Sample Design Data

| C1 | C2 | C3 | C4 | C5 | C6 | C7 |
|---|---|---|---|---|---|---|
| 871 | 1 | 1 | 10 | Low | -1 | 0 |
| 175 | 1 | 2 | 20 | Low | -1 | 0 |
| 484 | 1 | 2 | 10 | High | 1 | 0 |
| 332 | 1 | 1 | 20 | High | 1 | 0 |
| 146 | 1 | 2 | 10 | High | -1 | 1 |
| 571 | 1 | 1 | 20 | High | -1 | 1 |
| 878 | 1 | 1 | 10 | Low | 1 | 1 |
| 649 | 1 | 2 | 20 | Low | 1 | 1 |
| 953 | 2 | 2 | 10 | Low | -1 | 0 |
| 127 | 2 | 1 | 20 | Low | -1 | 0 |
| 533 | 2 | 1 | 10 | High | 1 | 0 |
| 966 | 2 | 2 | 20 | High | 1 | 0 |
| 404 | 2 | 1 | 10 | High | -1 | 1 |
| 817 | 2 | 2 | 20 | High | -1 | 1 |
| 588 | 2 | 2 | 10 | Low | 1 | 1 |
| 680 | 2 | 1 | 20 | Low | 1 | 1 |
| 870 | 3 | 2 | 10 | High | -1 | 0 |
| 838 | 3 | 1 | 20 | High | -1 | 0 |
| 511 | 3 | 1 | 10 | Low | 1 | 0 |
| 202 | 3 | 2 | 20 | Low | 1 | 0 |
| 686 | 3 | 1 | 10 | Low | -1 | 1 |
| 503 | 3 | 2 | 20 | Low | -1 | 1 |
| 427 | 3 | 2 | 10 | High | 1 | 1 |
| 968 | 3 | 1 | 20 | High | 1 | 1 |
| 99 | 4 | 1 | 10 | High | -1 | 0 |
| 381 | 4 | 2 | 20 | High | -1 | 0 |
| 695 | 4 | 2 | 10 | Low | 1 | 0 |
| 711 | 4 | 1 | 20 | Low | 1 | 0 |
| 476 | 4 | 2 | 10 | Low | -1 | 1 |
| 683 | 4 | 1 | 20 | Low | -1 | 1 |
| 740 | 4 | 1 | 10 | High | 1 | 1 |
| 972 | 4 | 2 | 20 | High | 1 | 1 |

Notice that the simulated response data is placed in variable C1, C2 contains the four block indices, and variables C3 through C7 contain the generated design values.

You would now proceed with your experiment, obtain the real response values, and analyze the data using one of the analysis of variance routines or the Analysis of Two-Level Designs procedure.

# Chapter 261

# Fractional Factorial Designs

## Introduction

This program generates two-level fractional-factorial designs of up to sixteen factors with blocking. Reports show the aliasing pattern that is used. The design rows may be output in standard or random order.

When generating a design, the program first checks to see if the design is among those listed on page 410 of Box and Hunter (1978). These designs are especially good. If the requested design is not listed in the above book, the design pattern is determined using the standard procedure in which the highest-order interactions are confounded first, and so on. The program makes certain that main effects are not aliased with each other.

An introduction to experimental design is presented in Chapter 83 on Two-Level Factorial Designs and will not be repeated here.

## Procedure Options

This section describes the options available in this procedure.

### Design Tab

This panel specifies the parameters that will be used to create the design values.

#### Data Storage Variables

##### Simulated Response Variable

This optional variable will contain a computer-simulated response value for each row. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

##### Block Variable

The variable to contain the block identification numbers. The blocks are numbered from one to B, where B is the number of blocks. This variable is optional. If this option is left blank, no blocks will be generated.

### First Factor Variable

This is where the group of variables that is to contain your design begins. The K-1 variables after this variable are also filled with data. The number of variables used is determined by the number of Factor Values boxes that contain data.

Warning: The program fills these variables with data, so any previous data will be lost.

## Data Storage Variables – Storage Options

### Sort Order

The order of the generated rows. The rows may be in random or standard order.

- **Random**

  The rows are randomly ordered (random blocks and random rows within blocks). Use this option when the order of application to experimental units is governed by the row number.

- **Standard**

  The rows are not randomly ordered. Instead, they are placed in standard order. Use this option when you want to quickly see the structure of the design.

## Experimental Setup

### Runs

The desired size (number of rows) of the experiment. This number must be a power of two. This number determines what fraction of a complete replicate is run. For example, suppose you are contemplating an experiment with seven factors and have budget for sixteen runs. A full replication would take $2^7 = 128$ runs. Hence, this design is a 1/8th rep (note that 16/128 = 1/8).

### Block Size

The number of experiments (runs) per block. This determines the number of blocks. This number must be a power of 2 (2, 4, 8, 16, etc.). Of course, the block size must be less than or equal to one half the number of runs.

### Factor Values

Each factor has two possible values (levels) which are specified here. These are the values that will be written to the database. The first value is used to represent the low value. The second value represents the high value. You may use both text and numeric values.

The number of variables created depends on how many of these boxes have values in them.

# Reports Tab

These options designate the variables to contain the design and the values that will be placed in those variables.

## Select Reports

### Design Info Report

Specifies whether to display this report.

### Aliases Report

Specifies whether to display this report.

## Report Options

### Aliases

One of the reports shows the confounding pattern among the columns of the design. However, when several factors are confounded, the number of terms aliased with each other gets huge. This option lets you limit the amount of information that the program displays.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Fractional Factorial Design

This section presents an example of how to generate an experimental design using this program. **CAUTION: since the purpose of this routine is to generate data, any existing data will be replaced. For this reason, you should begin with an empty database.**

In this example, we will show you how to generate a six-factor design using sixteen runs separated in blocks of four runs each. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Fractional Factorial Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Fractional Factorial Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Fractional Factorial Designs**. The Fractional Factorial Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Fractional Factorial Designs window, select the **Design tab**.
- Enter **1** in the **Simulated Response Variable** box.
- Enter **2** in the **Block Variable** box.
- Enter **3** in the **First Factor Variable** box.
- Select **16** in the **Runs** box.
- Select **4** in the **Block Size** box.
- Set **six of the Factor Values boxes** equal to **-1, 1**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## 1/4 Rep of a Six-Factor Design in Blocks of 4 Runs

| C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|---|---|---|---|---|---|---|---|
| 30.7 | 1 | -1 | -1 | 1 | 1 | 1 | -1 |
| 52.4 | 1 | 1 | 1 | -1 | -1 | -1 | 1 |
| 32.8 | 1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 73.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 42.6 | 2 | -1 | 1 | -1 | -1 | 1 | 1 |
| 7.9 | 2 | -1 | 1 | 1 | 1 | -1 | 1 |
| 65.9 | 2 | 1 | -1 | -1 | -1 | 1 | -1 |
| 70.6 | 2 | 1 | -1 | 1 | 1 | -1 | -1 |
| 10.6 | 3 | -1 | -1 | 1 | -1 | 1 | 1 |
| 96.6 | 3 | -1 | -1 | -1 | 1 | -1 | 1 |
| 51.1 | 3 | 1 | 1 | 1 | -1 | 1 | -1 |
| 15.6 | 3 | 1 | 1 | -1 | 1 | -1 | -1 |
| 16.5 | 4 | 1 | -1 | -1 | 1 | 1 | 1 |
| 53.8 | 4 | -1 | 1 | 1 | -1 | -1 | -1 |
| 5.4 | 4 | -1 | 1 | -1 | 1 | 1 | -1 |
| 60.6 | 4 | 1 | -1 | 1 | -1 | -1 | 1 |

Notice that the simulated response data is placed in variable C1, C2 contains the four block indices, and variables C3 through C8 contain the generated design values.

Note that since we selected the *random* order, your data will not appear in the same order as this example.

You would now proceed with your experiment, obtain the real response values, and analyze the data using one of the analysis of variance programs or the Analysis of Two-Level Design program.

## Design Information Section

**Design Information Section**

**Design:**
1/4 replication of 6 factors in 4 blocks of 4 experiments.

**Defining Contrast:**
i = ABCE = BCDF = ADEF

**Design Construction:**
Generate a reduced model of the factors [ A B C D ].
The remaining factors are aliased with interactions
of this reduced model as follows:
E = ABC
F = BCD

**Blocking Section**

**Block:**
Blocks were generated by confounding them with the
following interactions from the reduced model:
ABCD, CD

This report provides technical information about the design that was generated.

## Aliases Section

**One-Factor Aliases Section**

A+BCE+ABCDF+DEF
B+ACE+CDF+ABDEF
C+ABE+BDF+ACDEF
D+ABCDE+BCF+AEF
E+ABC+BCDEF+ADF
F+ABCEF+BCD+ADE

**Two-Factor Interaction Aliases Section**
AB+CE+ACDF+BDEF
AC+BE+ABDF+CDEF
AD+BCDE+ABCF+EF
AE+BC+ABCDEF+DF
AF+BCEF+ABCD+DE
BC+AE+DF+ABCDEF
BD+ACDE+CF+ABEF
BE+AC+CDEF+ABDF
BF+ACEF+CD+ABDE
CD+ABDE+BF+ACEF
CE+AB+BDEF+ACDF
CF+ABEF+BD+ACDE
DE+ABCD+BCEF+AF
DF+ABCDEF+BC+AE
EF+ABCF+BCDE+AD

This report lists the aliases of the main effects and low-order interactions. The number of aliases listed is controlled by the Aliases Shown option. This report provides technical information about the design that was generated.

From the first line of the report, we find that factor A (factor 1) is confounded with interactions BCE, DEF, and ABCDF. If any of the three-factor interactions are known to be real, this design would not be useful.

Note that no two-factor interactions (like AB or CD) are aliased with the main effects.

## Chapter 262

# Balanced Incomplete Block Designs

## Introduction

This module generates balanced incomplete block designs. Designs for up to ten treatments are available.

In order to make precise measurements of treatment means, uniform experimental conditions should be maintained when comparing a number of treatments. This insures that differences among the treatment means result from the application of the treatment and not from some extraneous factor. To achieve this, experimental trials are often grouped together into blocks. In such designs, conditions are kept constant within the blocks and allowed to vary between the blocks. The best known design of this type is the *randomized block* design. In this design, all treatments are present in each block.

Occasionally, the size of convenient blocks will not accommodate all the treatments of interest. For example, suppose you wanted to test four types of automobile tires for wear. An obvious choice for a block would be an automobile. You might select ten automobiles for the study. Assuming that the tires were rotated among the four positions, this experiment would control for differences in tire wear due to the type of automobile and the terrain that each traveled. However, what would you do if you wanted to test six types of tires. You could redesign the automobile, or you could adopt a *balanced incomplete block* design.

In a balanced incomplete block design, the treatments are assigned to the blocks so that every pair of treatments occurs together in a block the same number of times. This achieves the *balance* that is described in the title of the procedure. The balance means that all differences between treatments are measured with equal precision.

Following is an example of how four treatments are assigned to blocks with a natural size of three experimental units. Four blocks are required for this balanced incomplete block design.

| Block | Treatment |
|---|---|
| 1 | A B C |
| 2 | A B D |
| 3 | A C D |
| 4 | B C D |

Note that each treatment occurs three times in this experimental layout. Also note that each pair of treatments occurs twice. These are the basic properties of the balanced incomplete designs.

Box, Hunter, and Hunter (1978) point out the following rules when using such designs.

1. Randomly assign the numbers to the blocks.
2. Randomly assign the letters to the treatments.
3. Randomly assign the treatments within the blocks.
4. Randomly group blocks as replicates. A replicate is a complete set of all treatments.

If you take these steps, this design can be used effectively in those situations in which the block size and the number of treatments do not match.

# Design Limits

These designs were taken from Cochran and Cox (1992). We have included designs with up to ten treatments. The following table shows what block sizes are available for each number of treatments.

| Number of Treatments | Block Sizes Available |
|---|---|
| 4 | 2, 3 |
| 5 | 2, 3, 4 |
| 6 | 2, 3, 4, 5 |
| 7 | 2, 3, 4, 6 |
| 8 | 2, 4, 7 |
| 9 | 2, 4, 5, 6, 8 |
| 10 | 2, 3, 4, 5, 6, 9 |

Note that some block sizes are not available for certain numbers of treatments.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Store Trial Response In

This optional variable will contain a computer-simulated response value for each row that is generated. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

### Store First Factor In

The block identification numbers of each row of the design are stored in this variable. The treatment identification numbers (or letters) are stored in the variable immediately to the right.

Warning: The program fills these variables with data, so any previous data will be lost.

## Experimental Setup

### Block Size

This option contains the size of the blocks. That is, this is the number of experimental units that are contained in each block.

### Treatment Values

The values used to represent the treatments are specified here. These values may be letters, digits, words, or numbers. The list is delimited by blanks or commas. The number of treatments is implied by the number of items in this list.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Balanced Incomplete Block Design

This section presents an example of how to generate a balanced incomplete block design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

In this example, we will show you how to generate a design with four treatments in blocks of two experimental units each. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Balanced Incomplete Block Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Balanced Incomplete Block Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Balanced Incomplete Block Designs**. The Balanced Incomplete Block Designs procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Balanced Incomplete Block Designs window, select the **Design tab**.
- Set **Block Size** to **2**.
- Set **Treatment Values** to **1 2 3 4**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## BIBD with Four Treatments in Blocks of Two

| C1 | C2 | C3 |
|---|---|---|
| 150 | 1 | 1 |
| 980 | 1 | 2 |
| 893 | 2 | 3 |
| 378 | 2 | 4 |
| 940 | 3 | 1 |
| 140 | 3 | 3 |
| 72 | 4 | 2 |
| 116 | 4 | 4 |
| 154 | 5 | 1 |
| 469 | 5 | 4 |
| 418 | 6 | 2 |
| 324 | 6 | 3 |

Three variables are filled with data. The first variable contains the random response variable. The numbers in this column are random. Yours will not match those displayed here. The second variable, C2, contains the block identification number. The third variable, C3, contains the treatment number.

We note that this design calls for six blocks of two experimental units each.

To use this design, you would follow the randomization rules discussed earlier to obtain your experimental layout. After running your experiment, you would replace the random values in C1 with those obtained from your experiment. You would then analyze the data using the GLM procedure following the instructions for the randomized block design. You would specify blocks (C2) as Random and treatment (C3) as Fixed. The response variable would be C1. On the Model window of the GLM ANOVA procedure, you would set Which Model Terms to 'Up to 1-Way.' This forces the program to treat the block-by-treatment interaction as the error term.

## Analysis of Variance Table

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (C2) | 5 | No | S(AB) | S+bsA |
| B (C3) | 3 | Yes | S(AB) | S+asB |
| S(AB) | 3 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (C2) | 5 | 489979.6 | 97995.91 | 0.41 | 0.820890 | |
| B (C3) | 3 | 132675.3 | 44225.08 | 0.18 | 0.900707 | 0.057179 |
| S | 3 | 719385.8 | 239795.3 | | | |
| Total (Adjusted) | 11 | 1256687 | | | | |
| Total | 12 | | | | | |

* Term significant at alpha = 0.05

Since you are using random numbers for the response, the values of the sum of squares, mean squares, and F-ratios will not match those displayed here. However, the number of degrees of freedom will match.

Also note that the Expected Mean Square values are generated for a complete model. Since the balanced incomplete model is not complete, these values are incorrect.

# Chapter 263

# Latin Square Designs

## Introduction

This module generates Latin Square and Graeco-Latin Square designs. Designs for from three to ten treatments are available.

Latin Square designs are similar to randomized block designs, except that instead of the removal of one blocking variable, these designs are carefully constructed to allow the removal of two blocking factors. They accomplish this while reducing the number of experimental units needed to conduct the experiment.

Following is an example of a four treatment Latin Square. The experimental layout is as follows:

| | Columns | | | |
|---|---|---|---|---|
| **Rows** | **Col1** | **Col2** | **Col3** | **Col4** |
| Row 1 | A | B | C | D |
| Row 2 | B | C | D | A |
| Row 3 | C | D | A | B |
| Row 4 | D | A | B | C |

In the above table, the four treatments are represented by the four letters: A, B, C, and D. The letters are arranged so that each letter occurs only once within each row and each column. Notice that a simple random design would require 4 x 4 x 4 = 64 experimental units. This Latin Square needs only 16 experimental units—a reduction of 75%!

The influence of a fourth factor may also be removed from the design by introducing a second set of letters, this time lower case. This design is known as the *Graeco-Latin Square*.

| | Columns | | | |
|---|---|---|---|---|
| **Rows** | **Col1** | **Col2** | **Col3** | **Col4** |
| Row 1 | Aa | Bb | Cc | Dd |
| Row 2 | Bd | Ca | Db | Ac |
| Row 3 | Cb | Dc | Ad | Ba |
| Row 4 | Dc | Ad | Ba | Cb |

Four factors at four levels each would normally require 256 experimental units, but this design only requires 16—a reduction in experimental units of almost 94%!

The Graeco-Latin Square is formed by combining to orthogonal Latin Squares. Graeco-Latin Squares are available for all numbers of treatments except six.

## Latin Square Assumptions

It is important to understand the assumptions that are made when using the Latin Square design. The large reduction in the number of experimental units needed by this design occurs because it assumptions the magnitudes of the interaction terms are small enough that they may be ignored. That is, the Latin Square design is a main effects only design. Another way of saying this is that the treatments, the row factor, and the column factor affect the response independently of one another.

Assuming that there are no interactions is quite restrictive, so before you use this design you should be able to defend this assumption. In practice, the influence of the interactions is averaged into the experimental error of the analysis of variance table. We say that the experimental error is inflated. This results in a reduced F-ratio for testing the treatment factor, and a reduced F-ratio lessens the possibility of achieving statistical significance.

## Randomization

Probability statements made during the analysis of the experimental data require strict attention to the randomization process. The randomization process is as follows:

1. Randomly select a design from the set of orthogonal designs available.
2. Randomly assign levels of the row factor to the rows.
3. Randomly assign levels of the column factor to the columns.
4. Randomly assign treatments to the treatment letters (or numbers as the case may be).

## Orthogonal Sets

These designs were taken from Rao, Mitra, and Matthai (1966). We have included designs with up to ten treatments. The number of available squares depends on the number of treatments. The following table shows the number of orthogonal squares stored within this procedure.

| Number of Treatments | Number of Orthogonal Designs |
|---|---|
| 3 | 2 |
| 4 | 3 |
| 5 | 4 |
| 6 | 1 |
| 7 | 6 |
| 8 | 7 |
| 9 | 8 |
| 10 | 2 |

Graeco-Latin Squares are generated by combining two of the available orthogonal squares. Note that there are no six-level Graeco-Latin Squares.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Store Trial Response In

This optional variable will contain a computer-simulated response value for each row that is generated. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

#### Store First Factor In

The row values are stored in this variable. The column values are stored in the variable immediately to the right. The treatment values are stored in the variable immediately to the right of the column variable. If specified, the values of the second treatment are stored in the variable immediately to the right of the first treatment variable.

Warning: The program fills these variables with data, so any previous data will be replaced.

### Experimental Setup

#### Row Values

The values used to represent the rows are specified here. These values may be letters, digits, words, or numbers. The list is delimited by blanks or commas. The number of rows is implied by the number of items in this list. The number of row, column, and treatment values must be equal. From three to ten values are allowed.

#### Column Values

The values used to represent the columns are specified here. These values may be letters, digits, words, or numbers. The list is delimited by blanks or commas. The number of rows is implied by the number of items in this list. The number of row, column, and treatment values must be equal. From three to ten values are allowed.

#### Treatment 1 Values

The values used to represent the treatments are specified here. These values may be letters, digits, words, or numbers. The list is delimited by blanks or commas. The number of rows is implied by the number of items in this list. The number of row, column, and treatment values must be equal. From three to ten values are allowed.

#### Treatment 2 Values

The values used to represent the second set of treatments are specified here. These values may be letters, digits, words, or numbers. The list is delimited by commas. The number of rows is implied by the number of items in this list. The number of row, column, and treatment values must be equal. From three to ten values are allowed.

Note that this value is left blank unless you want to generate a Graeco-Latin Square.

## Experimental Setup – Orthogonal Designs

### Orthogonal Design Number I

Select one of the available orthogonal designs. The number of available orthogonal designs is given in the table in Orthogonal Sets section above. Good scientific protocol requires that you randomly choose which of these designs is used.

### Orthogonal Design Number II

This option is only used when the Treatment 2 Values box is non-blank (when you are generating a Graeco-Latin Square). Select a second of the available orthogonal designs to be combined with the first in forming a Graeco-Latin Square. The value here must be different from the value specified in Orthogonal Design I. Good scientific protocol requires that you randomly choose which of these designs is used.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Latin Square Design

This section presents an example of how to generate a Latin Square design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should begin with an empty database.**

In this example, we will show you how to generate a design with four treatments. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Latin Square Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Latin Square Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Latin Square Designs**. The Latin Square Designs procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Latin Square Designs window, select the **Design tab**.
- Set **Row Values** to **R1 R2 R3 R4**.
- Set **Column Values** to **C1 C2 C3 C4**.
- Set **Treatment 1 Values** to **A B C D**.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Four-Level Latin Square Design

| C1 | C2 | C3 | C4 |
|---|---|---|---|
| 996 | R1 | C1 | A |
| 838 | R1 | C2 | B |
| 134 | R1 | C3 | C |
| 216 | R1 | C4 | D |
| 747 | R2 | C1 | B |
| 754 | R2 | C2 | A |
| 121 | R2 | C3 | D |
| 295 | R2 | C4 | C |
| 641 | R3 | C1 | C |
| 936 | R3 | C2 | D |
| 237 | R3 | C3 | A |
| 208 | R3 | C4 | B |
| 362 | R4 | C1 | D |
| 639 | R4 | C2 | C |
| 781 | R4 | C3 | B |
| 876 | R4 | C4 | A |

Four variables are filled with data. The first variable contains the random response variable. The numbers in this column were selected at random. Yours will not match those displayed here. The second variable, C2, contains the row value. The third variable, C3, contains the column value. The fourth variable, C4, contains the treatment letter.

To use this design, you would follow the randomization rules discussed earlier to obtain your experimental layout. After running your experiment, you would replace the random values in C1 with those obtained from your experiment. You would then analyze the data using the GLM procedure. You would specify Factor 1 (C2) as Fixed (or Random as the case may be), Factor 2 (C3) as Fixed (or Random as the case may be), and treatment (C4) as Fixed. The response variable would be C1.

On the Model window of the GLM ANOVA procedure, you would set Which Model Terms to 'Up to 1-Way.' This forces the program to combine all interaction terms into an error term. The results will be similar to this.

## Analysis of Variance Table

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (C2) | 3 | Yes | S(ABC) | S+bcsA |
| B (C3) | 3 | Yes | S(ABC) | S+acsB |
| C (C4) | 3 | Yes | S(ABC) | S+absC |
| S(ABC) | 6 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (C2) | 3 | 80425.69 | 26808.56 | 0.33 | 0.806088 | 0.071380 |
| B (C3) | 3 | 614617.2 | 204872.4 | 2.50 | 0.156199 | 0.239479 |
| C (C4) | 3 | 284915.2 | 94971.73 | 1.16 | 0.399286 | 0.131679 |
| S | 6 | 491094.4 | 81849.06 | | | |
| Total (Adjusted) | 15 | 1471053 | | | | |
| Total | 16 | | | | | |

* Term significant at alpha = 0.05

Since you are using random numbers for the response, the values of the sum of squares, mean squares, and F-ratios will not match those displayed here. However, the number of degrees of freedom will match.

Note that only six degrees of freedom are available for the error term (S). This is a severe limitation of a Latin Square design with only four-levels. Often, you would replicate the experiment to obtain more error degrees of freedom.

Also note that the Expected Mean Square values are generated from the complete model assumption. Since the Latin Square is not complete (does not include all row-by-column-by-treatment combinations), these values are incorrect. The actual expected mean squares in this case would be S+4A, S+4B, and S+4C, respectively.

# Chapter 264

# Response Surface Designs

## Introduction

Response-surface designs are the only designs provided that allow for more than two levels. There are two general types of response-surface designs. The central-composite designs give five levels to each factor. The Box-Behnken designs give three levels to each factor.

The Central-Composite designs build upon the two-level factorial designs by adding a few center points and star points. A factor's five values are: *-a*, -1, 0, 1, and *a*. The value of *a* is determined by the number of factors in such a way that the resulting design is orthogonal. For example, if you are going to use either four or five factors, the value of *a* is 2.00.

The actual values of the levels are determined from these five values as follows:

1. The low-level value is assigned to -1.
2. The high-level value is assigned to 1.
3. The average of these two values is assigned to 0.
4. The values of *-a* and *a* are used to find the minimum and the maximum values.

For example, suppose we entered 50 for the low-level and 60 for the high level. Further, suppose there were four factors in the experiment. The levels would be

| Coded Level | Actual Level |
|---|---|
| ***-a*** | 45 |
| -1 | 50 |
| 0 | 55 |
| 1 | 60 |
| ***a*** | 65 |

The values of *a* depend on the number of factors in the design:

| Factors | Value of *a* |
|---|---|
| 2 | 1.41 |
| 3 | 1.73 |
| 4 | 2.00 |
| 5 | 2.00 |
| 6 | 2.24 |

The Box-Behnken designs have two differences from the central-composite designs. First, they usually use fewer runs. Second, they only use three levels while the central-composite designs use five.

The actual values of the levels are determined in the same manner as the central-composite designs, except that the value of *a* is ignored.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Simulated Response Variable

This optional variable will contain a computer-simulated response value for each row. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 100.

#### Block Variable

The variable to contain the block identification numbers. The blocks are numbered from one to B, where B is the number of blocks. This variable is optional. If this option is left blank, no blocks will be generated.

#### First Factor Variable

This is where the group of variables that is to contain your design begins. The K-1 variables after this variable are also filled with data. The number of variables used is determined by the number of Factor Values boxes that contain data. Up to six variables may be used.

Warning: The program fills these variables with data, so any previous data will be lost.

### Data Storage Variables – Storage Options

#### Sort Order

The order of the generated rows. The rows may be in random or standard order.

- **Random**

  The rows are randomly ordered (random blocks and random rows within blocks). Use this option when the order of application to experimental units is governed by the row number.

- **Standard**

  The rows are not randomly ordered. Instead, they are placed in standard order. Use this option when you want to quickly see the structure of the design.

## Experimental Setup

### Design Type

Specify whether to generate a *central-composite* or a *Box-Behnken* design. This selection controls the number of runs generated as well as the block size (if a blocking variable is present).

## Experimental Setup – Factor Values

### Factor Values

Each factor has three or five possible values (levels). The values associated with -1 and 1 are entered here.

If a Box-Behnken design was selected, the resulting three values will be -1,0,1. For example, if you entered 10 20 here, the resulting values would be 10, 15, and 20.

If a central-composite design was selected, the resulting five values will be -*a*, -1, 0, 1, *a*. For example, if you had four factors and entered 50 60 here, the resulting values would be 45, 50, 55, 60, and 65.

These are the values that will be written to the database. You can only use numeric values.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Response Surface Design

This section presents an example of how to generate an experimental design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

In this example, we will show you how to generate a three-factor central composite design with blocks. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Response Surface Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select New.
- Click the **Ok** button.

**2 Open the Response Surface Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Response Surface Designs**. The Response Surface Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Response Surface Designs window, select the **Design** tab.
- Enter 1 in the Simulated Response Variable box.
- Enter **2** in the **Block Variable** box.
- Enter 3 in the First Factor Variable box.
- Select **Standard** in the **Sort Order** list box.
- Select Central-Composite in the Design Type list box.
- Set three of the Factor Values boxes equal to -1 1.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Three-Factor Response-Surface Design

| C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|
| 415 | 1 | -1 | -1 | -1 |
| 767 | 1 | 1 | -1 | -1 |
| 905 | 1 | -1 | 1 | -1 |
| 848 | 1 | 1 | 1 | -1 |
| 135 | 1 | -1 | -1 | 1 |
| 366 | 1 | 1 | -1 | 1 |
| 80 | 1 | -1 | 1 | 1 |
| 655 | 1 | 1 | 1 | 1 |
| 970 | 1 | 0 | 0 | 0 |
| 343 | 1 | 0 | 0 | 0 |
| 955 | 1 | 0 | 0 | 0 |
| 865 | 2 | -1.73 | 0 | 0 |
| 85 | 2 | 1.73 | 0 | 0 |
| 15 | 2 | 0 | -1.73 | 0 |
| 140 | 2 | 0 | 1.73 | 0 |
| 964 | 2 | 0 | 0 | -1.73 |
| 84 | 2 | 0 | 0 | 1.73 |
| 384 | 2 | 0 | 0 | 0 |
| 188 | 2 | 0 | 0 | 0 |
| 599 | 2 | 0 | 0 | 0 |

Note that there are three replicates of the center points in each block. Note the star points represented by -1.73 and 1.73.

# Chapter 265

# Screening Designs

## Introduction

Screening designs are used to find the important factors from a large number (up to 31) of two-level factors. When the number of runs is 4, 8, 16, or 32 (powers of 2), the design is a regular fractional replication. When the number of runs is 12, 20, 24, or 28, the design used is a Plackett-Burman design.

This program uses the screening designs given in Lawson (1987). These designs make it possible to evaluate each main effect, although these are aliased with several interactions.

When you analyze the data from these designs, it is simplest to use our Multiple Regression routine. The Analysis of Two-Level Designs program can be used to analyze designs in which the number of runs is a power of 2 (the non-Plackett Burman designs).

## Procedure Options

This section describes the options available in this procedure.

### Design Tab

This panel specifies the parameters that will be used to create the design values.

#### Data Storage Variables

##### Simulated Response Variable

This optional variable will contain a computer-simulated response value for each row that is generated. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

##### First Factor Variable

This is where the group of variables that is to contain your design begins. The K-1 variables after this variable are also filled with data. The number of variables generated depends on the number of Factor Value boxes that contain data.

Warning: The program fills these variables with data, so any previous data will be lost.

## Data Storage Variables – Storage Options

### Sort Order

The order of the generated rows. The rows may be in random or standard order.

## Experimental Setup

### Runs

The desired size (number of rows) of the experiment. This number must be 4, 8, 12, 16, 20, 24, 28, or 32. This number determines which design is generated.

- **Random**

  The rows are randomly ordered (random blocks and random rows within blocks). Use this option when the order of application to experimental units is governed by the row number.

- **Standard**

  The rows are not randomly ordered. Instead, they are placed in standard order. Use this option when you want to quickly see the structure of the design.

## Experimental Setup – Factor Values

### Factor Values

Each factor has two possible values (levels), which are specified here. These are the values that will be written to the database. The first value is used to represent the low value. The second value represents the high value. You may use both text and numeric values, although we recommend that you stick with numeric values since these may be used in the regression program.

Enter a pair of values separated by a blank or comma, such as ‘-1 1’ or ‘0 1.’

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id’s

A list of the Template Id’s of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Screening Design

This section presents an example of how to generate an experimental design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

In this example, we will show you how to generate a six-factor design using 16 runs. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Screening Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Screening Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Screening Designs**. The Screening Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Screening Designs window, select the **Design tab**.
- Enter 1 in the Simulated Response Variable box.
- Enter 2 in the First Factor Variable box.
- Select Standard in the Sort Order list box.
- Set six of the Factor Values boxes equal to -1 1.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Six-Factor Screening Design in Sixteen Runs

| C1 | C2 | C3 | C4 | C5 | C6 | C7 |
|---|---|---|---|---|---|---|
| 653 | -1 | -1 | -1 | -1 | -1 | -1 |
| 730 | 1 | -1 | -1 | -1 | 1 | 1 |
| 867 | -1 | 1 | -1 | -1 | 1 | 1 |
| 785 | 1 | 1 | -1 | -1 | -1 | -1 |
| 732 | -1 | -1 | 1 | -1 | 1 | -1 |
| 725 | 1 | -1 | 1 | -1 | -1 | 1 |
| 433 | -1 | 1 | 1 | -1 | -1 | 1 |
| 690 | 1 | 1 | 1 | -1 | 1 | -1 |
| 459 | -1 | -1 | -1 | 1 | -1 | 1 |
| 470 | 1 | -1 | -1 | 1 | 1 | -1 |
| 672 | -1 | 1 | -1 | 1 | 1 | -1 |
| 798 | 1 | 1 | -1 | 1 | -1 | 1 |
| 909 | -1 | -1 | 1 | 1 | 1 | 1 |
| 49 | 1 | -1 | 1 | 1 | -1 | -1 |
| 398 | -1 | 1 | 1 | 1 | -1 | -1 |
| 481 | 1 | 1 | 1 | 1 | 1 | 1 |

Usually, you would specify the number of runs as close to the number of variables as possible, while still leaving some degrees of freedom for an estimate of error.

# Chapter 266

# Taguchi Designs

## Introduction

Taguchi experimental designs, often called orthogonal arrays (OA's), consist of a set of fractional factorial designs which ignore interaction and concentrate on main effect estimation. This program module generates the most popular set of Taguchi designs.

Taguchi uses the following convention for naming the orthogonal arrays: L*a*(*b*^*c*) where *a* is the number of experimental runs, *b* is the number of levels of each factor, and *c* is the number of variables. Designs can have factors with several levels, although two and three level designs are the most common. The $L_{18}$ design is perhaps the most popular.

When a design is generated, the levels of each factor are stored in the current database--replacing any data that is already there. No output reports are generated by this procedure.

## Procedure Options

This section describes the options available in this procedure.

### Design Tab

This panel specifies the parameters that will be used to create the design values.

#### Data Storage Variables

##### Simulated Response Variable

This optional variable will contain a computer-simulated response value for each row that is generated. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

##### First Factor Variable

This is where the group of variables that is to contain your design begins. The K-1 variables after this variable are also filled with data, where K is the number of variables specified.

Warning: The program fills these variables with data, so any previous data will be lost.

## Experimental Setup

### Design Type

This option designates the particular design that is to be generated. The available choices are:

- **L4 2^3**
  This design consists of up to 3 factors at 2 levels each. There are 4 rows.
- **L8 2^7**
  This design consists of up to 7 factors at 2 levels each. There are 8 rows.
- **L12 2^11**
  This design consists of up to 11 factors at 2 levels each. There are 12 rows.
- **L16 2^15**
  This design consists of up to 15 factors at 2 levels each. There are 16 rows.
- **L32 2^31**
  This design consists of up to 31 factors at 2 levels each. There are 32 rows.
- **L64 2^63**
  This design consists of up to 63 factors at 2 levels each. There are 64 rows.
- **L9 3^4**
  This design consists of up to 4 factors at 3 levels each. There are 9 rows.
- **L27 3^13**
  This design consists of up to 13 factors at 3 levels each. There are 27 rows.
- **L27' 3^22**
  This design consists of up to 22 factors at 3 levels each. There are 27 rows.
- **L16' 4^5**
  This design consists of up to 5 factors at 4 levels each. There are 16 rows.
- **L25 5^6**
  This design consists of up to 6 factors at 5 levels each. There are 25 rows.
- **L18 2^1 x 3^7**
  This design consists of one factor at 2 levels and up to 7 factors at 3 levels each. There are 18 rows.
- **L36 2^3 x 3^13**
  This design consists of up to 3 factors at 2 levels and up to 13 factors at 3 levels each. There are 36 rows.
- **L36' 2^11 x 3^12**
  This design consists of up to 11 factors at 2 levels and up to 12 factors at 3 levels each. There are 36 rows.

- **L54 2^1 x 3^25**

  This design consists of one factor at 2 levels and up to 25 factors at 3 levels each. There are 54 rows.

- **L32' 2^1 x 4^9**

  This design consists of one factor at 2 levels and up to 9 factors at 4 levels each. There are 32 rows.

- **L50 2^1 x 5^11**

  This design consists of one factor at 2 levels and up to 11 factors at 5 levels each. There are 50 rows.

## Experimental Setup – Factor Specification

### 2 Level Factors…5 Level Factors

The number of variables of this type (number of levels) that are generated. For example, if you selected L36 2^3 x 3^13 as the Design Type, you could specify up to three two-level factors and up to thirteen three-level factors. You would enter the number of two-level factors in the 2-Level Factors box and the number of three-level factors in the 3-Level Factors box. Entries in the unused boxes (such as 4-Level and 5-Level in this example) are ignored. If you ask for more than the maximum allowed, the maximum will be used.

Warning: The program fills these variables with data, so previous data may be lost.

# Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

## Specify the Template File Name

### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Taguchi Design

This section presents an example of how to generate an experimental design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

In this example, we will show you how to generate an L18 design. You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Taguchi Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Taguchi Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Taguchi Designs**. The Taguchi Designs procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Taguchi Designs window, select the **Design** tab.
- Enter **1** in the **Simulated Response Variable** box.
- Enter **2** in the **First Factor Variable** box.
- Select **L18  2^1 x 3^7** in the **Design Type** list box.
- Enter **1** in the **2-Level Factors** box.
- Enter **7** in the **3-Level Factors** box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Taguchi L18 Design

| C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
|---|---|---|---|---|---|---|---|---|
| 705 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 533 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 |
| 579 | 1 | 1 | 3 | 3 | 3 | 3 | 3 | 3 |
| 289 | 1 | 2 | 1 | 1 | 2 | 2 | 3 | 3 |
| 301 | 1 | 2 | 2 | 2 | 3 | 3 | 1 | 1 |
| 774 | 1 | 2 | 3 | 3 | 1 | 1 | 2 | 2 |
| 14 | 1 | 3 | 1 | 2 | 1 | 3 | 2 | 3 |
| 760 | 1 | 3 | 2 | 3 | 2 | 1 | 3 | 1 |
| 814 | 1 | 3 | 3 | 1 | 3 | 2 | 1 | 2 |
| 709 | 2 | 1 | 1 | 3 | 3 | 2 | 2 | 1 |
| 45 | 2 | 1 | 2 | 1 | 1 | 3 | 3 | 2 |
| 414 | 2 | 1 | 3 | 2 | 2 | 1 | 1 | 3 |
| 862 | 2 | 2 | 1 | 2 | 3 | 1 | 3 | 2 |
| 790 | 2 | 2 | 2 | 3 | 1 | 2 | 1 | 3 |
| 373 | 2 | 2 | 3 | 1 | 2 | 3 | 2 | 1 |
| 961 | 2 | 3 | 1 | 3 | 2 | 3 | 1 | 2 |
| 871 | 2 | 3 | 2 | 1 | 3 | 1 | 2 | 3 |
| 56 | 2 | 3 | 3 | 2 | 1 | 2 | 3 | 1 |

This shows the data that were generated in the dataset. You can use the Find/Replace facility of the spreadsheet if you want to change the values from 1, 2, 3 to something more meaningful.

# Chapter 267

# D-Optimal Designs

## Introduction

This procedure generates D-optimal designs  for multi-factor experiments with both quantitative and qualitative factors. The factors can have a mixed number of levels. Hence, you could use this procedure to design an experiment with two quantitative factors having three levels each and a qualitative factor having seven levels.

D-optimal designs are constructed to minimize the generalized variance of the estimated regression coefficients. In the multiple regression setting, the matrix **X** is often used to represent the data matrix of independent variables. D-optimal designs minimize the overall variance of the estimated regression coefficients by maximizing the determinant of **X'X**. Designs that are D-optimal have been shown to be nearly optimal for several other criterion that have been proposed as well.

When would you use D-optimal designs? When you have a limited budget and cannot run a completely replicated factorial design. For example, suppose you want to study the response to three factors: A with three levels, B with four levels, and C with eight levels. One complete replication of this experiment would require 3 x 4 x 8 = 96 points (we use the word 'point' to mean an experimental unit). Suppose you can afford only 20 points. Which 20 of the 96 possible should you use? The D-optimal design algorithm provides a reasonable choice.

## D-Optimal Design Overview

This section provides a brief overview of how the D-optimal design algorithm works. It will provide a general understanding of what the algorithm is trying to accomplish so that you can make intelligent choices for the various options.

Suppose you are studying the influence of height and weight on blood pressure. If you believe that a linear (straight line) relationship exists, you will only need to look at two height values and two weight values. An experiment designed to study this relationship would require four treatment combinations. However, if you decide that the relationship may be curvilinear, you will have to include at least three levels for each factor which results in nine treatment combinations. Clearly, the appropriate experimental design depends on the anticipated functional relationship between the response variable and the factors of interest.

The D-optimal algorithm works as follows. First, specify an approximate mathematical model which defines the functional form of the relationship between the response (Y) and the independent variables (the factors). Next, generate a set of possible candidate points based on this model. Finally, from these candidates select the subset that maximizes the determinant of the **X'X** matrix. This is the D-optimal design. The details of this algorithm are given in Atkinson and Donev (1992).

The number of possible designs grows rapidly as the complexity of the model increases. This number is usually so large that an exhaustive search of all possible designs for a given sample size is not feasible.

The D-optimal algorithm begins with a randomly selected set of points. Points in and out of the current design are exchanged until no exchange can be found that increases the determinant of **X'X**. To cut down on the running time, the number of points considered during any one iteration may be limited.

Unfortunately, this method does not guarantee that the global maximum is found. To overcome this, the algorithm is repeated several times in hopes that at least one iteration leads to the global maximum. For this reason 50 or 100 random starting sets are needed. (During the testing of the algorithm, we found that some designs required 500 starts to obtain the global maximum.)

## Factor Scaling

This algorithm deals with both *quantitative* (continuous) and *qualitative* (discrete) factors. The levels of quantitative factors are scaled so that the minimum value is -1 and the maximum value is 1. Qualitative factors are included as a set of variables. For example, suppose that a qualitative variable has four values. Three independent variables are created to represent this factor:

| Original | X1 | X2 | X3 |
|---|---|---|---|
| 1 | -10 | 0 | 0 |
| 2 | 0 | -1 | 0 |
| 3 | 0 | 0 | -1 |
| 4 | 1 | 1 | 1 |

As you can see, each of these variables compares a separate group with the last group. Also note that the number of generated variables is always one less than the number of levels.

## Duplicates (Replicates)

The measurement of experimental error is extremely important in the analysis of an experiment. In most cases, if an estimate of experimental error is not available, the data from the experiment cannot be analyzed. One of the best estimates of experiment error comes from points that are duplicates (often called replicates) of each other. Since D-optimal designs are often used in situations with limited budgets, the experimenter is often tempted to ignore the need for duplicates and instead add points with additional treatment combinations. The tenth commandments for experimental design should be "Thou shalt have at least four duplicates in an experiment."

Unfortunately, the D-optimal design algorithm ignores the need for duplicates. Instead, you have to add them after the experimental design has be found. So what you do is set aside at least four points from the algorithm. For example, suppose you have budget for 20 design points. You would tell the program that you have only 16 points. The algorithm would find the best 16 point design. You would then duplicate four of the resulting design points to provide an estimate of experimental error. We recommend that you spread these duplicates out across the experiment so you can have some indication as to whether the magnitude of the experimental error is constant across all treatment settings.

## Specifying a Model

Selecting an appropriate model is subjective by nature. Often, you will know very little about the true functional form of the relationship between the response and the factor variables. A common approach is to assume that a second-order Taylor-series approximation will work fairly well. You are assuming that the true function may be approximated by parabolic surface in the neighborhood of interest. Cutting down on the complexity of the model reduces the number of points that must be added to the experimental design.

When dealing with qualitative factors, you generally limit the model to first order interactions. Higher order interactions may be studied later when a complete experiment can be run.

## Augmenting an Existing Design

Occasionally, you will want to add more points to an existing experimental design. This may be accomplished by forcing the algorithm to include points that are read from the spreadsheet. The D-optimal algorithm will pick the most useful additional points from the list of candidate points. One of the attractive features of the D-optimal design algorithm is that you can refine the model as your knowledge of it increases.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Simulated Response Variable

This optional variable will contain a random response value for each row in the design. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design will provide the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

#### First Factor Variable

If the Input Data Type is set to Factor Values, the final design is stored in a set of contiguous columns of the spreadsheet, beginning with this column. Be careful not to overwrite existing data. If you have four factors, the design will be stored in this variable and the next three to the right. Existing data will be lost!

If the Input Data Type is set to Expanded Matrix, an index is stored in this variable that represents whether the row is used in the design. If the row is not in the optimum design, a zero is stored. If the row is in the optimum design, the number of times it occurs is stored here.

### First Expanded Variable

This option specifies the first variable in which to store the expanded version of the selected design. The rest of the expanded design variables will be stored in the variables to the right. Use this option if you want to output the expanded design matrix for use in the multiple regression procedure.

Warning: The program fills these variables with data, so existing data will be replaced.

## Data Storage Variables – Storage Options

### Rename Factor Variables with Factor Labels

The names of the factors that were used in the model statement are used to rename the variables in which the design is stored.

### Clear Existing Data

Clear all existing data in the design variables before writing the new design data. This is especially useful if you are experimenting with several designs of different sizes. You will not be warned that data is being lost. The data will be cleared and the new design written automatically.

## Experimental Setup

### N Per Block

This option specifies the required sample size. If you are not using blocks, enter a single number giving the total sample size. The sample size must be large enough to fit the designated model. If it is not large enough, you will be shown the minimum number of points necessary.

If you are using blocks, enter the sample size for each block, separated by blanks or commas. These sample sizes do not have to be equal, although they usually are. For example, if you have three blocks, you might enter 8,8,12 which would give an overall sample size of 28. The first block will have 8 points, the second 8 points, and the third 12 points.

You must be careful when specifying blocks when you also have forced design points. In this case, the first few blocks are matched with the forced design points. The size of the blocks must match the number of forced points. For example, suppose you have already run two blocks of four each and you want to augment this with three blocks of six each. You would have eight forced points. The entry in this field would be 4,4,6,6,6,6. If you entered 4,3,7,6,6 an error would occur because the forced points cannot be assigned exactly to one or more blocks. The bottom line is, you cannot force partial blocks into the design.

### Input Variables (Candidate and Forced)

When specified, these variables contain either a set of points to be forced into the final design, a set of candidate points from which the design is to be selected, or both. The data must be arranged so that the forced points are located at the top of the spreadsheet followed by any candidate points. When candidate points are specified, no additional candidate points are generated. If you want to force points in the design and choose the rest from among those generated by the model statement, the total number of rows in these variables must equal the total number of forced rows specified below.

Note that these variables are matched with the factors specified in the model after those factors have been sorted.

Qualitative factors must be entered using positive integers (1, 2, 3, etc.). You cannot use any other identifiers. If you have data entered using some other scheme (such as A, B, C, etc.), you will have to recode the values so that they are positive integers.

Quantitative factors must be scaled so that the minimum value is -1 and the maximum value is 1. For example, suppose an existing design has a factor whose values are 10, 15, and 20. Here the minimum is 10 and the maximum is 20. You would transform these using the formula

**Scaled = (Original + Original - Max - Min) / (Max - Min)**

Since, in this example, Max = 20 and Min = 10, the transformation reduces to New = (Original + Original - 30)/10 = Original /5 - 3. You would create a new variable using the transformation Original /5-3. This transformation would give 10/5 - 3 = -1, 15/5 - 3 = 0, and 20/5 - 3 = 1. That is, the new variable would contain -1's, 0's, and 1's instead of 10's, 15's, and 20's.

### Number Duplicates

It is very important to have duplicates of at least some of the design points to provide an estimate of experimental error. This option designates the number of duplicates to be generated. The first design point is duplicated, then the second, and so on. Even though this option is convenient, we recommend that you pick appropriate points for duplication by looking at scatter plots of the design.

If your design includes blocking, you should not create duplicates since that will give erroneous block sizes. Rather, you should manually create duplicates.

### Input Data Type

If you have Input Variables specified, this option specifies the type of data contained in those variables. Two types of data are possible.

- **Factor Values**

  Specifies that the input data contains indices of each factor. An expanded design matrix will be generated from these factor indices using the designated model. This is the more common data type.

- **Expanded Matrix**

  Specifies that the input dataset contains the expanded design matrix. That is, the quadratic, cubic, and interaction terms have been created. The model statement is not used. You would use this option when you want to specify the candidate design set in more detail than is allowed by the program. The expanded matrix must include the intercept (a column of one's) if one is to be included in the model.

### Forced Points

The number of rows in the Input Variables that should be forced into the final design. These rows must be located at the top of the database, before any candidate points. If the number of forced points is equal to the number of points read in, the generated design matrix is used. Otherwise, the additional rows are used as candidate points and no other rows are generated.

### Optimize the Design for this Model

Your design is optimized for the model specified here. Specify main effects (factors) with names consisting of one or more letters, such as A B C. Specify interactions using an asterisk (*), such as A*B. You can use the bar (|) symbol (see examples below) as a shorthand method to specify a complete model. You can use parentheses. You can separate terms with blanks or the '+' (plus)

sign. Duplicate terms are removed during the evaluation of the model. Note that the main effects are always sorted in alphabetical order.

Some examples will help to indicate how the model syntax works:

| | | |
|---|---|---|
| C + B + A + B*A + C*A | = | A+B+C+A*B+A*C (Note the sorting!) |
| A\|B | = | A+B+A*B |
| B\|A | = | A+B+A*B |
| A\|B A*A B*B | = | A+B+A*B+A*A+B*B |
| A\|A\|B\|B (Max Term Order=2) | = | A+B+A*B+A*A+B*B |
| A\|B\|C | = | A+B+C+A*B+A*C+B*C+A*B*C |
| (A+B)*(C+D) | = | A*C+A*D+B*C+B*D |
| (A+B)\|C | = | A+B+C+(A+B)*C |
| | = | A+B+C+A*C+B*C |

You can experiment with various expressions by viewing the Model Terms report.

For quantitative factors, each term represents a single variable in the expanded design matrix. For qualitative variables, each term represents a set of variables in the expanded design matrix.

Note that qualitative terms should not be squared or cubed. That is, if A is a qualitative factor, you would not include A*A or an A*A*A in your model.

### Max Term Order

This option specifies that maximum number of factors that can occur in an interaction term. For example, A*B*C is a third order interaction term and if this option were set to 2, the A*B*C would be removed from the model.

This option is particularly useful when used with the bar notation to remove unwanted terms.

### Qualitative Factors and Levels

List any qualitative factors here followed by the number of levels given in parenthesis. Factors in the model which do not appear here are assumed to be quantitative (continuous). For example, you might enter A(5),B(4),C(7) to indicate three qualitative factors, one with five levels, the next with four levels, and the third with seven levels. Of course, the names used here must match the names used in the model statement.

### Max Iterations

Specify the number of times the algorithm is started with a new random design. Often 50 or 100 iterations are necessary and 500 is not unheard of. As the number of Inclusion Points and Removal Points are increased (see below), the number of iterations may be decreased.

We suggest that you increase this value until the optimal design is found on several iterations as reported in the Determinant Analysis report.

### Inclusion Points

This is the number of candidate points considered for addition during an iteration. Instead of considering all candidate points, only this many are used. A value between 1 and *Nc*-1 (where *Nc* is the number of candidate points) may be used. Usually, a value near *Nc*/2 is adequate.

### Removal Points

This is the number of points currently in the design that will be considered for removal during a particular iteration. A value between 1 and *N* (the desired sample size) is used. Setting this value smaller than *N* speeds up the search, but reduces the possibility of finding the optimal design.

### Include Intercept

This option specifies whether to include the intercept in the expanded design matrix. Usually, the intercept is left out of mixture designs. The intercept is automatically deleted in designs with more than one block.

## Reports Tab

This panel specifies the reports that will be generated.

### Select Reports

#### Factor Report - Expanded Design Matrix Report

These options control which reports are displayed. Some of the reports may be fairly lengthy, so you will often want to omit them.

### Report Options

#### Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

#### Decimal Places

Specify the number of decimal places shown when displaying the design.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

### Select a Template to Load or Save

#### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

## Example 1 – D-Optimal Design with 10 Points, 3 Factors

This section presents an example of how to generate a D-optimal design using this program. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should begin with an empty database.**

In this example, we will show you how to generate a 10-point design for a study involving three quantitative factors. We want the design optimized to estimate a second-order response surface model.

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the D-Optimal Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the D-Optimal Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The D-Optimal Designs procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design and data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **Simulated Response Variable** to **1**.
- Set the **First Factor Variable** to **2**.
- Set the **First Expanded Variable** to **6**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data**.
- Set **Optimize the Design for this Model** to **A|A|B|B|C|C**.
- Set **Max Term Order** to **2**.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Check **Candidate Points Report**.
- Check **Expanded Design Matrix Report**.
- Set **Decimal Places** to **0**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## 10-Point, 3 Factor D-Optimal Design

| C1 | A | B | C | C5 Int't | Ax | Bx | Cx | A_A | A_B | A_C | B_B | B_C | C_C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 160 | -1 | -1 | -1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 878 | 1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 163 | 0 | 0 | -1 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 435 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | 1 | -1 | 1 |
| 926 | 1 | 1 | -1 | 1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | -1 | 1 |
| 466 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 988 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 431 | 1 | -1 | 1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 |
| 766 | -1 | 0 | 1 | 1 | -1 | 0 | 1 | 1 | 0 | -1 | 0 | 0 | 1 |
| 713 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Several variables in the spreadsheet are filled with data. The first variable (C1) contains the random response variable. Your values will not match those displayed here. The second, third, and fourth variables (A, B, and C) contain the actual design. You would replace the -1's with the corresponding factor's minimum value, the 1's with the maximum value, and the 0's with the average of the two.

The variables from Intercept to C_C contain the expanded design matrix. Each variable is generated by multiplying the appropriate factor values. For example, in the first row, A_B is found by multiplying the value for A, which is -1, by the value for B, which is also -1. The result is 1. The intercept is set to one for all rows. The expanded matrix is usually saved so that the design can be analyzed using multiple regression.

To use this design, you would randomly assign these ten points to the ten experimental units.

## Factor Section

| Name | Number Values | Type | Value1 | Value2 | Value3 |
|---|---|---|---|---|---|
| A | 3 | Quantitative | -1.0000 | 0.0000 | 1.0000 |
| B | 3 | Quantitative | -1.0000 | 0.0000 | 1.0000 |
| C | 3 | Quantitative | -1.0000 | 0.0000 | 1.0000 |

A total of 27 observations will be needed for one replication.

This report summarizes the factors that were included in the design. The last line of this report gives the number of observations required for one complete replication of the experiment. This value is the product of the number of levels for each factor.

### Name

The symbol(s) used to represent the factor.

### Number Values

The number of values (levels) generated for each factor. For qualitative factors, this value was set in the Qualitative Factors and Levels box of the Design panel. For quantitative factors, this value is one more that the highest exponent used with this term. For example, if the model includes an A*A and nothing of a higher order, this value will be three.

### Type

A factor is either quantitative or qualitative.

### Value1 - Value 3

These columns list the individual values that are used as the levels of each factor when generating the expanded design matrix based on the model. Notice that the smallest is always -1 and the largest is always 1.

When the expanded design matrix is input directly, these values should be ignored.

## Model Terms Section

| Variables Needed | Term |
|---|---|
| 1 | A |
| 1 | B |
| 1 | C |
| 1 | A*A |
| 1 | A*B |
| 1 | A*C |
| 1 | B*B |
| 1 | B*C |
| 1 | C*C |
| 9 | Model Total |

This report shows the terms generated by your model. You should check this report carefully to make sure that the generated model matches what you wanted. The last line of the report gives the total number of degrees of freedom (except for the intercept) required for your model. This number plus one is the minimum size of the D-optimal design for this model.

### Variables Needed

The number of degrees of freedom (expanded design variables) required for this term.

### Term

The name of each term.

## D-Optimal Design

| Original Row | Factors A | B | C |
|---|---|---|---|
| 1 | -1 | -1 | -1 |
| 3 | 1 | -1 | -1 |
| 5 | 0 | 0 | -1 |
| 7 | -1 | 1 | -1 |
| 9 | 1 | 1 | -1 |
| 13 | -1 | 0 | 0 |
| 17 | 0 | 1 | 0 |
| 20 | 0 | -1 | 1 |
| 25 | -1 | 1 | 1 |
| 27 | 1 | 1 | 1 |

This report gives the points in the D-optimal design.

### Original Row

This is the row number of the point from the list of candidate points. It is only useful in those cases in which you provided the list of candidate points manually.

### Factors (A B C)

These are the values of the factors. For example, the first row sets A, B, and C to -1. Remember that these are scaled values. You would transform them back into their original metric using the formula:

**Original = (Scaled(Max - Min) + Max + Min)/2**

For example, suppose the original metric for factor A is minimum = 10 and maximum =20. The original values would be calculated as follows:

| Scaled | Formula | Original |
|---|---|---|
| -1 | (-1(20-10)+20+10)/2 | 10 |
| 0 | (0(20-10)+20+10)/2 | 15 |
| 1 | (1(20-10)+20+10)/2 | 20 |

The values 10, 15, and 20 represent the three levels of factor A that are used in the design. They would replace the -1, 0, and 1 displayed in this report.

## Determinant Analysis Section

| Rank | Determinant of X'X | D-Efficiency | Percent of Maximum |
|---|---|---|---|
| 1 | 1327104 | 40.95 | 100.00 |
| 2 | 1327104 | 40.95 | 100.00 |
| 3 | 1048576 | 40.00 | 79.01 |
| 4 | 1048576 | 40.00 | 79.01 |
| 5 | 1048576 | 40.00 | 79.01 |
| 6 | 1048576 | 40.00 | 79.01 |
| 7 | 1048576 | 40.00 | 79.01 |
| 8 | 921600 | 39.49 | 69.44 |
| 9 | 921600 | 39.49 | 69.44 |
| 10 | 802816 | 38.95 | 60.49 |
| 11 | 802816 | 38.95 | 60.49 |
| 12 | 802816 | 38.95 | 60.49 |
| 13 | 802816 | 38.95 | 60.49 |
| 14 | 802816 | 38.95 | 60.49 |
| 15 | 802816 | 38.95 | 60.49 |
| 16 | 746496 | 38.66 | 56.25 |
| 17 | 589824 | 37.76 | 44.44 |
| 18 | 589824 | 37.76 | 44.44 |
| 19 | 589824 | 37.76 | 44.44 |
| 20 | 589824 | 37.76 | 44.44 |

The maximum was achieved on 2 of 30 iterations.

This report shows the largest twenty determinants. The main purpose of this report is to let you decide if enough iterations have been run so that a global maximum has been found. Unless the maximum value was achieved on at least five iterations, you should double the number of iterations and rerun the procedure.

In this example, the top value occurred on only two iterations. In practice we would probably try another 200 iterations to find out if this is the global maximum.

### Rank

Only the top twenty are shown on this report. The values are sorted by the determinant.

### Determinant of X'X

This is the value of the determinant of **X'X** which is the statistic that is being maximized. This value is sometimes called the generalized variance of the regression coefficients. Since this value occurs in the denominator of the variance of each regression coefficient, maximizing it has the effect of reducing the variance of the estimated regression coefficients.

### D-Efficiency

D-efficiency is the relative number of runs (expressed as a percent) required by a hypothetical orthogonal design to achieve the same determinant value. It provides a way of comparing designs across different sample sizes.

$$DE = 100\left(\frac{\left|X'X\right|^{1/p}}{N}\right)$$

where $p$ is the total number of degrees of freedom in the model and $N$ is the number of points in the design.

### Percent of Maximum

This is the percentage that the determinant on this row is of the best determinant found.

## Individual Degree of Freedom Section

| Number | Name | Diagonal of X'X | Diagonal of X'X Inv |
|---|---|---|---|
| 1 | Intercept | 10.0000 | 0.861111 |
| 2 | A | 7.0000 | 0.250000 |
| 3 | B | 8.0000 | 0.166667 |
| 4 | C | 8.0000 | 0.166667 |
| 5 | A*A | 7.0000 | 0.722222 |
| 6 | A*B | 6.0000 | 0.250000 |
| 7 | A*C | 6.0000 | 0.250000 |
| 8 | B*B | 8.0000 | 0.861111 |
| 9 | B*C | 7.0000 | 0.194444 |
| 10 | C*C | 8.0000 | 0.861111 |
| | | | |
| Determinant | | 1327104 | |
| D-Efficiency | | 40.95345 | |
| Trace | | 4.583333 | |
| A-Efficiency | | 21.81818 | |

This report shows the diagonal elements of the **X'X** and its inverse. Since the variance of each term is proportional to diagonal elements from the inverse of **X'X**, the last column of this report lets you compare those variances. From this report you can determine if the coefficients will be estimated with the relative precision that is desired.

For example, we can see from this example that them main effects will be estimated with the greatest precision—usually a desirable quality in a design.

### Number

An arbitrary sequence number.

### Name

The name of the term.

### Diagonal of X'X

The diagonal element of this term in the **X'X** matrix.

### Diagonal of X'X Inv

The diagonal element of this term in the **X'X** inverse matrix. See the discussion above for an understanding of how this value might be interpreted.

### Determinant

This is the value of the determinant of **X'X** which is the statistic that is being maximized. This value is sometimes called the generalized variance of the regression coefficients. Since this value occurs in the denominator of the variance of each regression coefficient, maximizing it has the effect of reducing the variance of the estimated regression coefficients.

### D-Efficiency

D-efficiency is the relative number of runs (expressed as a percent) required by a hypothetical orthogonal design to achieve the same determinant value. It provides a way of comparing designs across different sample sizes.

$$DE = 100\left(\frac{\left|X'X\right|^{1/p}}{N}\right)$$

where $p$ is the total number of degrees of freedom in the model and $N$ is the number of points in the design.

### Trace

This is the value of the trace of **X'X-inverse** which is associated with A-optimality.

### A-Efficiency

D-efficiency is the relative number of runs (expressed as a percent) required by a hypothetical orthogonal design to achieve the same trace value. It provides a way of comparing designs across different sample sizes.

$$AE = 100\left(\frac{p}{trace\left(N\left(X'X\right)^{-1}\right)}\right)$$

where $p$ is the total number of degrees of freedom in the model and $N$ is the number of points in the design.

## Candidate Points Section

| Original Row | Factors A | B | C |
|---|---|---|---|
| 1 | -1 | -1 | -1 |
| 2 | 0 | -1 | -1 |
| 3 | 1 | -1 | -1 |
| 4 | -1 | 0 | -1 |
| 5 | 0 | 0 | -1 |
| 6 | 1 | 0 | -1 |
| 7 | -1 | 1 | -1 |
| 8 | 0 | 1 | -1 |
| 9 | 1 | 1 | -1 |
| 10 | -1 | -1 | 0 |
| 11 | 0 | -1 | 0 |
| 12 | 1 | -1 | 0 |
| 13 | -1 | 0 | 0 |
| 14 | 0 | 0 | 0 |
| 15 | 1 | 0 | 0 |
| 16 | -1 | 1 | 0 |
| 17 | 0 | 1 | 0 |
| 18 | 1 | 1 | 0 |
| 19 | -1 | -1 | 1 |
| 20 | 0 | -1 | 1 |
| 21 | 1 | -1 | 1 |
| 22 | -1 | 0 | 1 |
| 23 | 0 | 0 | 1 |
| 24 | 1 | 0 | 1 |
| 25 | -1 | 1 | 1 |
| 26 | 0 | 1 | 1 |
| 27 | 1 | 1 | 1 |

This report gives a list of candidate points from which the D-optimal design points were selected.

### Original Row

This is an arbitrary identification number.

### Factors (A B C)

These are the values of the factors. For example, the first row sets A, B, and C to -1. Remember that these are scaled values. You would transform them back into their original metric using the formula:

**Original = (Scaled(Max - Min) + Max + Min)/2**

For example, suppose the original metric for factor A is minimum = 10 and maximum =20. The original values would be calculated as follows:

| Scaled | Formula | Original |
|---|---|---|
| -1 | (-1(20-10)+20+10)/2 | 10 |
| 0 | (0(20-10)+20+10)/2 | 15 |
| 1 | (1(20-10)+20+10)/2 | 20 |

The values 10, 15, and 20 represent the three levels of factor A. They would replace the -1, 0, and 1 displayed in this report.

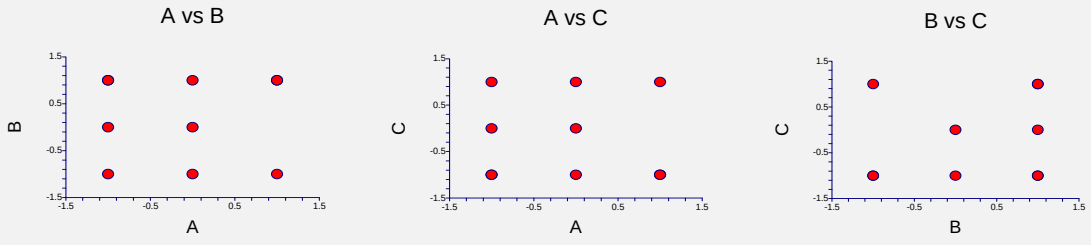
## Expanded Design Matrix Section

| Row | Variable Intercept | A | B | C | A*A | A*B | A*C | B*B | B*C | C*C |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 0 | -1 | -1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 3 | 1 | 1 | -1 | -1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 4 | 1 | -1 | 0 | -1 | 1 | 0 | 1 | 0 | 0 | 1 |
| 5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 6 | 1 | 1 | 0 | -1 | 1 | 0 | -1 | 0 | 0 | 1 |
| 7 | 1 | -1 | 1 | -1 | 1 | -1 | 1 | 1 | -1 | 1 |
| 8 | 1 | 0 | 1 | -1 | 0 | 0 | 0 | 1 | -1 | 1 |
| 9 | 1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | -1 | 1 |
| 10 | 1 | -1 | -1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| 11 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 12 | 1 | 1 | -1 | 0 | 1 | -1 | 0 | 1 | 0 | 0 |
| 13 | 1 | -1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 14 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 16 | 1 | -1 | 1 | 0 | 1 | -1 | 0 | 1 | 0 | 0 |
| 17 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 18 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| 19 | 1 | -1 | -1 | 1 | 1 | 1 | -1 | 1 | -1 | 1 |
| 20 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 1 | -1 | 1 |
| 21 | 1 | 1 | -1 | 1 | 1 | -1 | 1 | 1 | -1 | 1 |
| 22 | 1 | -1 | 0 | 1 | 1 | 0 | -1 | 0 | 0 | 1 |
| 23 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 24 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| 25 | 1 | -1 | 1 | 1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 26 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 27 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

This report gives a list of candidate points expanded so that each individual term may be seen. The report is useful to show you how the expanded matrix looks. Each variable is generated by multiplying the appropriate factor values. For example, in the first row, A_B is found by multiplying the value for A, which is -1, by the value for B, which is also -1. The result is 1. The intercept is set to one for all rows.

If you want to constrain the design space, you could cut and paste these values back into the spreadsheet and then eliminate points that cannot occur.

## Scatter Plots of Design




Finally, we ran the D-optimal design through the Scatter Plot procedure so that we could visually see how the design values are placed. It might be useful to create a 3-D scatter plot since we are dealing with three factors. Unfortunately, we have found that the 3-D plot is only useful interactively—motion is necessary to gain insights from the plot. Since this is not possible in the documentation, we suggest that you experiment with this on your own.

From these three scatter plots, we can see the configuration of the points fairly well. It appears that the B*C term is missing two points while the A*B and A*C terms are missing only one. Using this information, we would want to arrange our factors in such a way that the B*C term is the least likely to have an interaction.

## Example 2 – Two Factors

This section presents an example of how to generate and analyze a D-optimal design involving two factors. Suppose we want to study the effect of two factor variables, A and B, on a response variable, Y. A and B happen to be quantitative variables and there is reason to believe that a second-order response surface design will work well. A full replication of this design requires nine points. In addition, four more are required to provide an estimate of experimental error. However, we can only afford eight. We will create a D-optimal design with six of the experimental units and use the remaining two as duplicates to provide the estimate of experimental error.

We want to analyze the response surface for values of A between 10 and 20 and values of B between 1 and 3.

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the D-Optimal Designs window.

**1 Open a new (empty) dataset.**

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the D-Optimal Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design and data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **Simulated Response Variable** to **1**.
- Set the **First Factor Variable** to **2**.
- Set the **First Expanded Variable** to **5**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data**.
- Set **N Per Block** to **6**.
- Set **Optimize the Design for this Model** to **A|A|B|B**.
- Set **Max Term Order** to **2**.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Set **Decimal Places** to **0**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## 6-Point, 2 Factor D-Optimal Design

| C1 | A | B |
|---|---|---|
| 896 | -1 | -1 |
| 372 | 0 | -1 |
| 323 | 1 | -1 |
| 770 | 0 | 0 |
| 218 | -1 | 1 |
| 446 | 1 | 1 |

Variables A and B give the design. The Determinant Analysis Section showed that the maximum was achieved on 25 of the 30 iterations. Hence, we assume that the algorithm converged to the global maximum.

Next, we add the two duplicates to the design. When only a few duplicates are available, we like to have them in the middle, so we will duplicate the two rows having zero values. We choose random numbers for the two new response values. The resulting design appears as follows.

## 6-Point Design with Two Duplicates

| C1 | A | B |
|---|---|---|
| 896 | -1 | -1 |
| 372 | 0 | -1 |
| 374 | 0 | -1 |
| 323 | 1 | -1 |
| 770 | 0 | 0 |
| 774 | 0 | 0 |
| 218 | -1 | 1 |
| 446 | 1 | 1 |

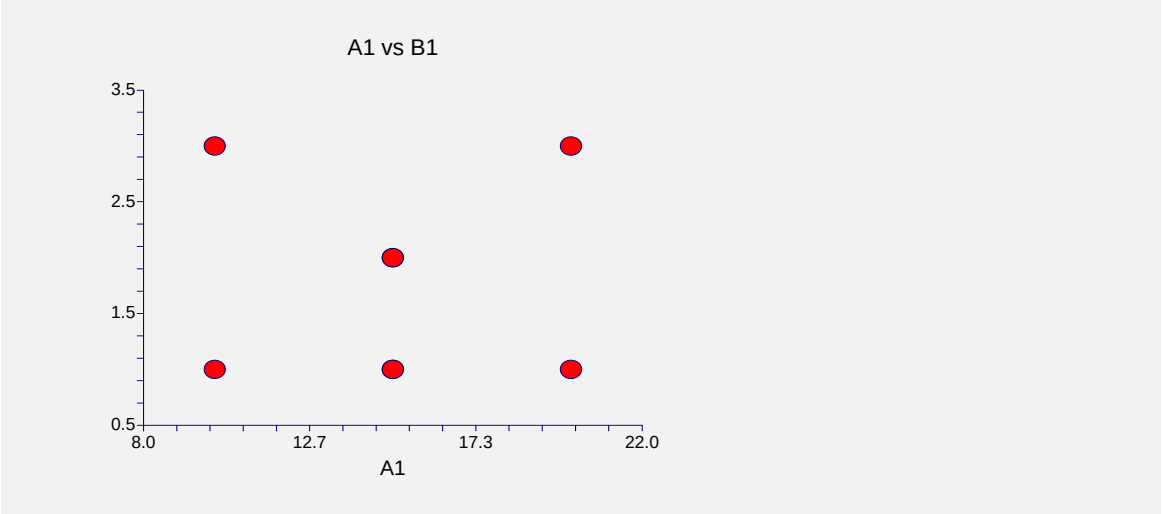
Next, we change the factor values back to their original scale. Factor A went from 10 to 20 and factor B went from 1 to 3. We call the two new variables A1 and B1. While we are at it, we also create other variables of the expanded design matrix. The resulting database appears as follows.

## 6-Point Design in Expanded Form

| C1 | A | B | A1 | B1 | A1_B1 | A1_A1 | B1_B1 |
|---|---|---|---|---|---|---|---|
| 896 | -1 | -1 | 10 | 1 | 10 | 100 | 1 |
| 372 | 0 | -1 | 15 | 1 | 15 | 225 | 1 |
| 374 | 0 | -1 | 15 | 1 | 15 | 225 | 1 |
| 323 | 1 | -1 | 20 | 1 | 20 | 400 | 1 |
| 770 | 0 | 0 | 15 | 2 | 30 | 225 | 4 |
| 774 | 0 | 0 | 15 | 2 | 30 | 225 | 4 |
| 218 | -1 | 1 | 10 | 3 | 30 | 100 | 9 |
| 446 | 1 | 1 | 20 | 3 | 60 | 400 | 9 |

We could continue this exercise by running these data through the multiple regression program and paying particular attention to the Multicollinearity Section and the Eigenvalues of Centered Correlations Section. When we did this, we found that multicollinearity seemed to be a problem in the original scale, but not in the -1 to 1 scale used by the D-optimal algorithm.

## Plot of Design




In order to better understand the design, we look at a scatter plot of the two factors. Remember that this began as a six-point design. We can see from this plot that the optimum configuration puts points at each corner and in the middle—just what we would expect. Viewing the design configuration is extremely important.

Remember that we duplicated the two center points of this design.

# Example 3 – Three Factors with Blocking

This section presents an example of how to generate and analyze a D-optimal design involving three factors with blocking.

Suppose we want to study the effect of three quantitative factor variables (A, B, and C) on a response variable. There is reason to believe that a second-order response surface design will work well. A full replication of this design requires twenty-seven experimental units. The manufacturing process that we are studying produces items in batches of four at a time. Because of this and the limited budget available for this study, we decide to use three batches (which we will call 'Blocks') of four points each.

You may follow along here by making the appropriate entries or load the completed template **Example3** from the Template tab of the D-Optimal Designs window.

### 1 Open a new (empty) dataset.

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

### 2 Open the D-Optimal Designs window.

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design and data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **Simulated Response Variable** to **1**.
- Set the **First Factor Variable** to **2**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data**.
- Set **N Per Block** to **4,4,4**.
- Set **Optimize the Design for this Model** to **A|B|C A*A B*B C*C**.
- Set **Max Term Order** to **2**.
- Set **Max Iterations** to **100**.
- Set **Inclusion Points** to **45**. This is approximately (3)(3)(3)(3)/2 which is the number of blocks times the product of the number of levels in each factor, all divided by two.
- Set **Removal Points** to **11**. This is one less than the total number of points desired.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Set **Decimal Places** to **0**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

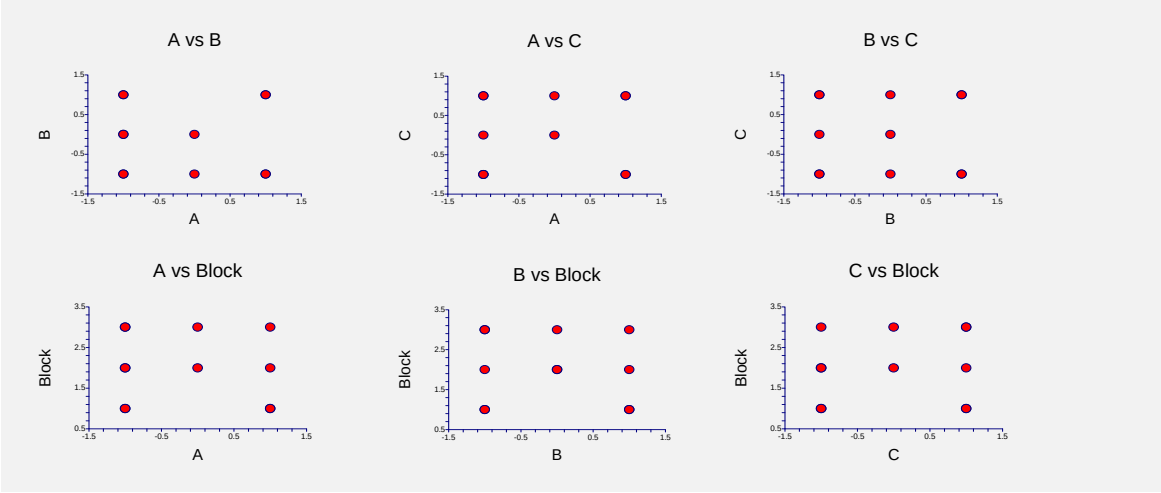
## 12-Point, 3 Factor D-Optimal Design with Blocking

| C1 | A | B | C | Blocks |
|---|---|---|---|---|
| 194 | -1 | -1 | -1 | 2 |
| 221 | 1 | -1 | -1 | 1 |
| 615 | -1 | 0 | -1 | 3 |
| 191 | -1 | 1 | -1 | 1 |
| 629 | 1 | 1 | -1 | 2 |
| 505 | 0 | -1 | 0 | 3 |
| 244 | -1 | 0 | 0 | 2 |
| 113 | -1 | -1 | 1 | 1 |
| 680 | 1 | -1 | 1 | 3 |
| 150 | 0 | 0 | 1 | 2 |
| 166 | -1 | 1 | 1 | 3 |
| 241 | 1 | 1 | 1 | |

Variables A, B, C, and Blocks give the design. The Determinant Analysis Section showed that the maximum was achieved on 12 of the 100 iterations. Hence, we assume that the algorithm converged to the global maximum.

In order to visually analyze the design, we generate the scatter plots for each pair of variables in the design.

## Plot of Design




We can see from these plots that each of the interactions seems to be well represented—only a few points are missing from each and none of these are on the corners. The design seems pretty good. We decide to use the interactions with blocks as the measure of experimental error, so no other duplicates are need.

As a exercise, try adding one more block to this experiment. You will notice that each of the two-way interaction plots are completely full.

# Example 4 – Adding Points to an Existing Design

This section presents an example of how to augment additional points to an existing design.

Suppose a standard three factor design has been run. Each factor has two levels. The design was blocked into two blocks of four points each. The design values are contained in the DOPT3.S0 database. This design allows only first-order (linear) terms to be fit.

Suppose that you wish to add more points to the design so that a second-order response surface may be fit. Specifically, suppose you want to add one more block of four points to extend the model from first to second order. What four points should be added?

You may follow along here by making the appropriate entries or load the completed template **Example4** from the Template tab of the D-Optimal Designs window.

### 1 Open DOPT3.S0.

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** directory.
- Select the file **DOPT3.S0**.
- Click the **Ok** button.

### 2 Open the D-Optimal Designs window.

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design and data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **First Factor Variable** to **5**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data**.
- Set **N Per Block** to **4,4,4**.
- Set **Input Variables (Candidate and Forced)** to **A-C**.
- Set **Forced Points** to **8**.
- Set **Optimize the Design for this Model** to **A|B|C A*A B*B C*C**.
- Set **Max Term Order** to **2**.
- Set **Max Iterations** to **30**.
- Set **Inclusion Points** to **5**.
- Set **Removal Points** to **5**.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Set **Decimal Places** to **0**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Augmented D-Optimal Design with Blocking

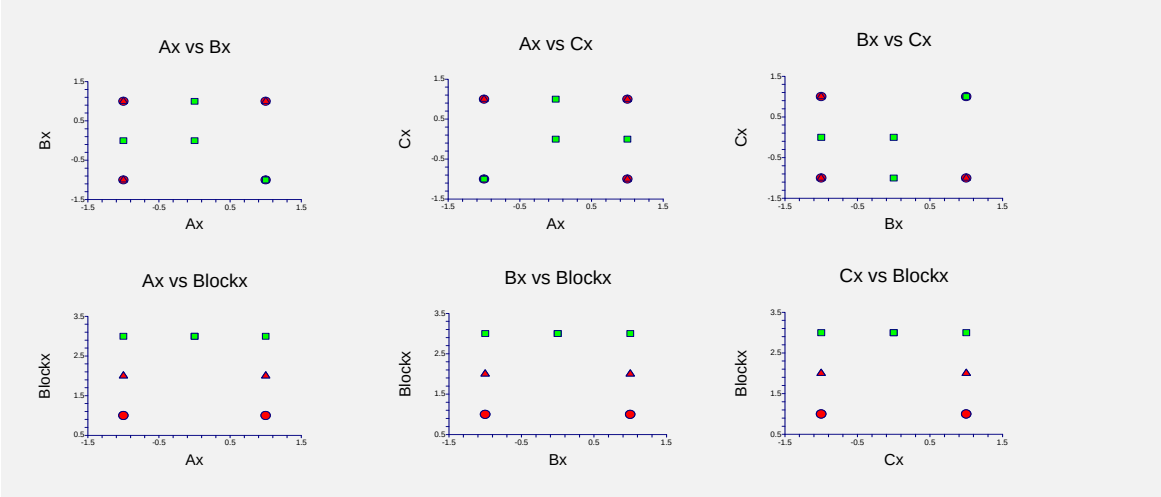
| A | B | C | Blocks |
|---|---|---|---|
| -1 | -1 | -1 | 1 |
| 1 | 1 | -1 | 1 |
| 1 | -1 | 1 | 1 |
| -1 | 1 | 1 | 1 |
| 1 | -1 | -1 | 2 |
| -1 | 1 | -1 | 2 |
| -1 | -1 | 1 | 2 |
| 1 | 1 | 1 | 2 |
| -1 | 0 | -1 | 3 |
| 0 | 1 | -1 | 3 |
| 1 | -1 | 0 | 3 |
| 0 | 0 | 0 | 3 |

Variables A, B, C, and Blocks give the design. The new block is shown as the last four rows of the design.

The Determinant Analysis Section showed that the maximum was achieved on 9 of the 30 iterations. Hence, we assume that the algorithm converged to the global maximum.

In order to visually analyze the design, we generate the scatter plots for each pair of variables in the design.

## Plot of Design



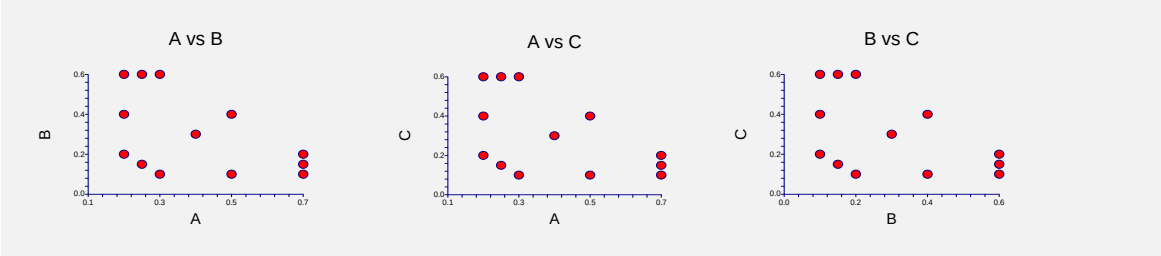

We set the plotting symbols in the scatter plots so that the new points are displayed as squares. It is interesting to see where these points were added.

## Example 5 – Mixture Design

This section presents an example of how to generate a mixture design. Mixture designs are useful in situations in which the factors are constrained to sum to a total. The interest is in the proportions of each factor, not the absolute amounts. For example, the proportions of the components of a chemical solution must sum to one.

Suppose that you wish to design a first-order mixture experiment for a chemical that has three components (which we will label as A, B, and C). In this case, you will not code the factor levels from -1 to 1. Rather, the factor levels will be coded from zero to one. Because of this constraint, the intercept will not be fit in this model.

In this particular case, we will constrain the design space by only entering certain points in the list of candidate points. The candidate points are contained in the database named DOPT_MIXED.S0. The following plots show the design space for each pair of factors. Remember that these factors are constrained so that the missing factor is equal to one minus the sum of the other two. Hence, if A is 0.7 and B is 0.2, then C must be 0.1.




The task for the algorithm is to pick the ten best points from the thirteen that are shown here.

You may follow along here by making the appropriate entries or load the completed template **Example5** from the Template tab of the D-Optimal Designs window.

**1 Open DOPT_MIXED.S0.**

- From the File menu of the NCSS Data window, select **Open**.
- Select the **Data** directory.
- Select the file **DOPT_MIXED.S0**.
- Click the **Ok** button.

**2 Open the D-Optimal Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design and data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **First Factor Variable** to **4**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data**.
- Set **N Per Block** to **10**.
- Set **Input Variables (Candidate and Forced)** to **1-3**.
- Set **Forced Points** to **0**.
- Set **Optimize the Design for this Model** to **A|B|C**.
- Set **Max Term Order** to **2**.
- Set **Max Iterations** to **30**.
- Set **Inclusion Points** to **5**.
- Set **Removal Points** to **5**.
- **Remove the check** from the **Include Intercept** check box.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Set **Decimal Places** to **4**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Mixture Design

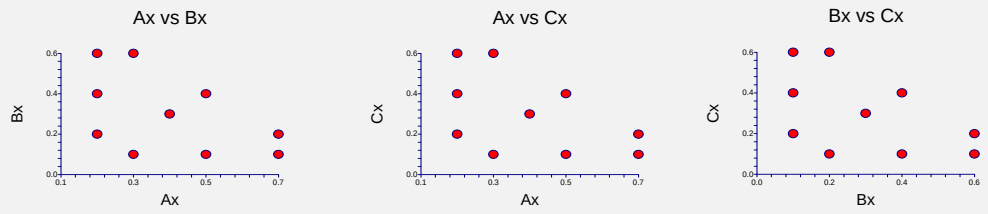
| Original Row | Factors A | B | C |
|---|---|---|---|
| 1 | 0.7000 | 0.1000 | 0.2000 |
| 2 | 0.2000 | 0.6000 | 0.2000 |
| 3 | 0.7000 | 0.2000 | 0.1000 |
| 4 | 0.2000 | 0.2000 | 0.6000 |
| 5 | 0.3000 | 0.6000 | 0.1000 |
| 6 | 0.3000 | 0.1000 | 0.6000 |
| 8 | 0.2000 | 0.4000 | 0.4000 |
| 9 | 0.5000 | 0.1000 | 0.4000 |
| 11 | 0.5000 | 0.4000 | 0.1000 |
| 13 | 0.4000 | 0.3000 | 0.3000 |

Columns A, B, and C give the design. The original row from the candidate list is shown as the first column of the report.

The Determinant Analysis Section showed that the maximum was achieved on 30 of the 30 iterations. Hence, we assume that the algorithm converged to the global maximum.

In order to visually analyze the design, we generate the scatter plots for each pair of variables in the design.

## Plot of Design




It is interesting to compare these plots with those produced earlier to see which points were kept by the algorithm.

# Example 6 – Qualitative Factors

This section presents an example of how to design an experiment with qualitative and quantitative factors.

Suppose your experimental situation involves two quantitative variables, A and B, and a qualitative variable C that has five possible levels. You want to fit a second-order response surface to the quantitative variables. Also, you want to fit all two-way interactions among these factors. You have budget for an 18-point design (you will add four duplicates later).

You may follow along here by making the appropriate entries or load the completed template **Example6** from the Template tab of the D-Optimal Designs window.

### 1 Open a new (empty) dataset.

- From the File menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the D-Optimal Designs window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **D-Optimal Designs**. The procedure window will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the data storage.**

- On the D-Optimal Designs window, select the **Design tab**.
- Set the **First Factor Variable** to **1**.
- Check **Rename Factor Variables with Factor Labels**.
- Check **Clear Existing Data.**
- Set **N Per Block** to **18**.
- Set **Optimize the Design for this Model** to **A|B|C A*A B*B**.
- Set **Max Term Order** to **2**.
- Set **Qualitative Factors and Levels** to **C(5)**.
- Set **Max Iterations** to **30**.
- Set **Inclusion Points** to **20**.
- Set **Removal Points** to **18**.
- **Remove the check** from the **Include Intercept** check box.

**4 Specify the reports.**

- On the D-Optimal Designs window, select the **Reports tab**.
- Set **Decimal Places** to **0**.

**5 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Design with Qualitative Factors

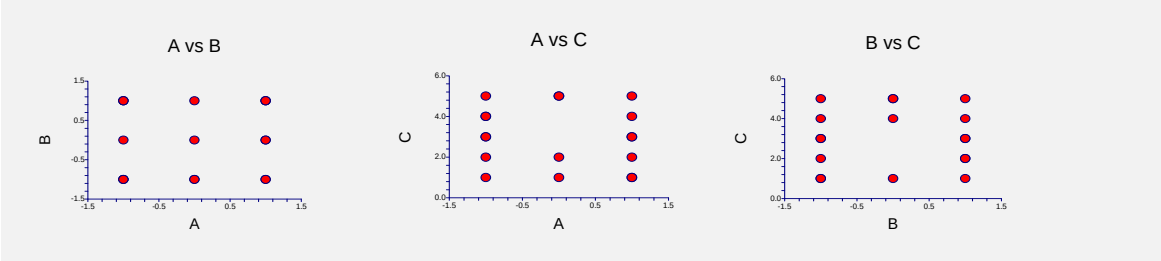
| Original | Factors | | |
|---|---|---|---|
| Row | A | B | C |
| 1 | -1 | -1 | 1 |
| 3 | 1 | -1 | 1 |
| 5 | 0 | 0 | 1 |
| 9 | 1 | 1 | 1 |
| 11 | 0 | -1 | 2 |
| 16 | -1 | 1 | 2 |
| 18 | 1 | 1 | 2 |
| 19 | -1 | -1 | 3 |
| 21 | 1 | -1 | 3 |
| 25 | -1 | 1 | 3 |
| 27 | 1 | 1 | 3 |
| 28 | -1 | -1 | 4 |
| 33 | 1 | 0 | 4 |
| 34 | -1 | 1 | 4 |
| 38 | 0 | -1 | 5 |
| 40 | -1 | 0 | 5 |
| 42 | 1 | 0 | 5 |
| 44 | 0 | 1 | 5 |

Columns A, B, and C give the design. Notice that column C simply gives the level for factor C—it was not rescaled. Also note that the levels of factor C are numbered arbitrarily. This means that only the pattern is important, not the particular level. For example, in this solution, there are only three level 2's and three level 4's. In the next solution, there might be three level 3's and three level 4's.

The Determinant Analysis Section showed that the maximum was achieved on 5 of the 30 iterations. Hence, we assume that the algorithm converged to the global maximum.

In order to visually analyze the design, we generate the scatter plots for each pair of variables in the design.

## Plot of Design




It is interesting to note that all nine positions were filled for the interaction of the two quantitative factors, A and B. However, some points were omitted for the AC interaction and the BC interaction.

# Chapter 268

# Design Generator

## Introduction

This program generates factorial, repeated measures, and split-plots designs with up to ten factors. The design is placed in the current database.

## Crossed Factors

Two factors are *crossed* if all levels of one factor occur with each level of the second factor. No distinction needs to be made as to whether a factor is random or fixed. Factorial and randomized block designs are examples of designs that contain crossed factors.

## Nested Factors

In the repeated measures and split-plot designs, at least one of the factors is nested in another factor. A factor is *nested* when all levels of this factor do not occur with each level of another factor. For example, suppose a study is being made to compare the heart rate of males and females. Five males and five females are selected. One factor in the study would be gender with two levels: male and female. Another factor would be individual with ten levels: P1, P2, …, and P10. Since five of the ten individuals are in the males group and the other five individuals are in the females group, individuals are nested within gender.

The basic structure of *repeated measures* and *split-plot* designs is identical. The difference between the two is in the way the factor levels are assigned within the individual factor. Consider an exercise study in which heart rate readings are to be made on an individual at five different points in time. If the amounts of exercise is assigned at random before each reading, the design is a split plot. If the amounts of exercise follow the same pattern for each individual, the design is a repeated measures.

# Procedure Options

This section describes the options available in this procedure.

## Design Tab

This panel specifies the parameters that will be used to create the design values.

### Data Storage Variables

#### Store Trial Response In

This optional variable will contain a computer-simulated response value for each row that is generated. These values may be used as the response variable in an analysis of variance or regression analysis to check that the analysis of this design provides the answers that you are looking for. The values themselves are from a uniform random-number generator that generates numbers between 0 and 1000.

#### Store First Factor In

The first factor is stored in this variable. Each additional factor that is specified is stored in the variables immediately to the right of this variable. A factor is specified when values are entered into its Factor Values box.

Warning: The program fills these variables with data, so any previous data will be replaced.

### Experimental Setup

#### Factor (1 to 12) Values

The values used to represent the rows are specified here. These values may be letters, digits, words, or numbers. The list is delimited by blanks or commas. The number of levels of a factor corresponds to the number of values that are listed here.

To specify a nested factor, use the word Nested followed by the number of levels within a group. For example, entering ‘Nested 4’ signifies a design in which four individuals are placed in each group. The number of groups is found by crossing the factors before the nested factor.

An easy way to replicate a design is to specify a nested factor as the last factor with the number of replicates specified as the number of levels.

## Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.

### Specify the Template File Name

#### File Name

Designate the name of the template file either to be loaded or stored.

## Select a Template to Load or Save

### Template Files

A list of previously stored template files for this procedure.

### Template Id's

A list of the Template Id's of the corresponding files. This id value is loaded in the box at the bottom of the panel.

# Example 1 – Three-by-Four Factorial Design with Three Replicates

This section presents an example of how to degenerate a three-by-four factorial design with three replicates per treatment combination. To run this example, take the following steps. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

You may follow along here by making the appropriate entries or load the completed template **Example1** from the Template tab of the Design Generator window.

**1 Open a new (empty) dataset.**

- From the **File** menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Design Generator window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Design Generator**. The Design Generator procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Design Generator window, select the **Design tab**.
- Enter **1 2 3** in the **Factor 1 Values (A)** box.
- Enter **1 2 3 4** in the **Factor 2 Values (B)** box.
- Enter **Nested 3** in the **Factor 3 Values (C)** box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Three-by-Four Design with Three Replicates

| C1 | C2 | C3 | C4 |
|---|---|---|---|
| 436 | 1 | 1 | 1 |
| 965 | 1 | 1 | 2 |
| 409 | 1 | 1 | 3 |
| 425 | 1 | 2 | 4 |
| 19 | 1 | 2 | 5 |
| 247 | 1 | 2 | 6 |
| 720 | 1 | 3 | 7 |
| 265 | 1 | 3 | 8 |
| 424 | 1 | 3 | 9 |
| 504 | 1 | 4 | 10 |
| 785 | 1 | 4 | 11 |
| 819 | 1 | 4 | 12 |
| 771 | 2 | 1 | 13 |
| 689 | 2 | 1 | 14 |
| 802 | 2 | 1 | 15 |
| 298 | 2 | 2 | 16 |
| 985 | 2 | 2 | 17 |
| 339 | 2 | 2 | 18 |
| 224 | 2 | 3 | 19 |
| 163 | 2 | 3 | 20 |
| 250 | 2 | 3 | 21 |
| 216 | 2 | 4 | 22 |
| 830 | 2 | 4 | 23 |
| 886 | 2 | 4 | 24 |
| 711 | 3 | 1 | 25 |
| 194 | 3 | 1 | 26 |
| 57 | 3 | 1 | 27 |
| 967 | 3 | 2 | 28 |
| 430 | 3 | 2 | 29 |
| 404 | 3 | 2 | 30 |
| 579 | 3 | 3 | 31 |
| 200 | 3 | 3 | 32 |
| 585 | 3 | 3 | 33 |
| 620 | 3 | 4 | 34 |
| 131 | 3 | 4 | 35 |
| 190 | 3 | 4 | 36 |

Notice that the simulated response is placed in variable C1, C2 contains the three values for factor 1, C2 contains the four values of factor 2, and C3 contains the value of the nested factor. When these data are analyzed, C3 will be ignored.

You would now proceed with your experiment, obtain the real response values, and analyze the data using either the Analysis of Variance procedure or the GLM procedure. The output will appear as follows.

## ANOVA for 3-by-4 Factorial

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (C2) | 2 | Yes | S(AB) | S+bsA |
| B (C3) | 3 | Yes | S(AB) | S+asB |
| AB | 6 | Yes | S(AB) | S+sAB |
| S(AB) | 24 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (C2) | 2 | 83609.72 | 41804.86 | 0.59 | 0.563864 | 0.113222 |
| B (C3) | 3 | 200053.6 | 66684.55 | 0.94 | 0.438592 | 0.181412 |
| AB | 6 | 832589.6 | 138764.9 | 1.95 | 0.113775 | 0.521955 |
| S | 24 | 1709695 | 71237.28 | | | |
| Total (Adjusted) | 35 | 2825948 | | | | |
| Total | 36 | | | | | |

* Term significant at alpha = 0.05

Of course, your F-Ratios will be different because you are using a different set of random numbers. However, you will be able to see the number of degrees of freedom that are associated with each factor.

# Example 2 – Randomized Block Design

This section presents an example of how to degenerate a randomized block design with three blocks and four treatments. To run this example, take the following steps. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

You may follow along here by making the appropriate entries or load the completed template **Example2** from the Template tab of the Design Generator window.

### 1 Open a new (empty) dataset.

- From the **File** menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

### 2 Open the Design Generator window.

- On the menus, select **Analysis**, then **Design of Experiments**, then **Design Generator**. The Design Generator procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

### 3 Specify the design parameters.

- On the Design Generator window, select the **Design tab**.
- Enter **1 2 3** in the **Factor 1 Values (A)** box.
- Enter **A B C D** in the **Factor 2 Values (B)** box.
- Make sure that the **Factor 3 Values (C)** box is blank.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Randomized Block Design

| C1 | C2 | C3 |
|---|---|---|
| 319 | 1 | A |
| 622 | 1 | B |
| 768 | 1 | C |
| 600 | 1 | D |
| 985 | 2 | A |
| 126 | 2 | B |
| 997 | 2 | C |
| 943 | 2 | D |
| 119 | 3 | A |
| 481 | 3 | B |
| 911 | 3 | C |
| 514 | 3 | D |

Notice that the simulated response is placed in variable C1, C2 contains the three values for the blocks, and C3 contains the value of the treatment.

It is important to remember that when you use this design, you must randomly assign treatments to the four letters and randomly assign the physical blocks to the three block numbers.

You would now proceed with your experiment, obtain the real response values, and analyze the data using the GLM procedure. You would specify blocks (C2) as random and treatment (C3) as fixed. You would set the Which Model Terms option of the Model tab to Up to 1-Way. After running the analysis, the output appears as follows.

## ANOVA for Randomized Block Design

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (C2) | 2 | No | S(AB) | S+bsA |
| B (C3) | 3 | Yes | S(AB) | S+asB |
| S(AB) | 6 | No | | S |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (C2) | 2 | 140324.7 | 70162.34 | 0.79 | 0.495026 | |
| B (C3) | 3 | 431012.9 | 143671 | 1.62 | 0.280672 | 0.211531 |
| S | 6 | 531277.3 | 88546.22 | | | |
| Total (Adjusted) | 11 | 1102615 | | | | |
| Total | 12 | | | | | |

* Term significant at alpha = 0.05

# Example 3 – Repeated Measures Design

This section presents an example of how to degenerate a repeated measures design with three groups, two individuals per group, and two treatments which we will label 'Pre' and 'Post.' To run this example, take the following steps. **CAUTION: since the purpose of this routine is to generate (not analyze) data, you should always begin with an empty database.**

You may follow along here by making the appropriate entries or load the completed template **Example3** from the Template tab of the Design Generator window.

**1 Open a new (empty) dataset.**

- From the **File** menu of the NCSS Data window, select **New**.
- Click the **Ok** button.

**2 Open the Design Generator window.**

- On the menus, select **Analysis**, then **Design of Experiments**, then **Design Generator**. The Design Generator procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

**3 Specify the design parameters.**

- On the Design Generator window, select the **Design tab**.
- Enter **1 2 3** in the **Factor 1 Values (A)** box.
- Enter **Nested 2** in the **Factor 2 Values (B)** box.
- Enter **Pre Post** in the **Factor 3 Values (C)** box.

**4 Run the procedure.**

- From the Run menu, select **Run Procedure**. Alternatively, just click the Run button (the left-most button on the button bar at the top).

## Repeated Measures Design

| C1 | C2 | C3 | C4 |
|---|---|---|---|
| 560 | 1 | 1 | Pre |
| 701 | 1 | 1 | Post |
| 874 | 1 | 2 | Pre |
| 5 | 1 | 2 | Post |
| 353 | 2 | 3 | Pre |
| 26 | 2 | 3 | Post |
| 43 | 2 | 4 | Pre |
| 319 | 2 | 4 | Post |
| 196 | 3 | 5 | Pre |
| 390 | 3 | 5 | Post |
| 520 | 3 | 6 | Pre |
| 433 | 3 | 6 | Post |

Notice that the simulated response is placed in variable C1. Variable C2 contains the three group values which are sometimes referred to as the *Between* factor. Variable C3 contains the identification numbers of the six individuals required for this design. Notice that each individual is placed in only one group (C2). Variable C4 contains the pre-post labels. The design is ready for analysis by the GLM procedure.

You would now proceed with your experiment, obtain the real response values, and analyze the data using the GLM procedure. You would specify variable (C2) as fixed, variable (C3), as nested, and variable C4 as fixed. After running the analysis, the output appears as follows.

## ANOVA for Repeated Measures Design

**Expected Mean Squares Section**

| Source Term | DF | Term Fixed? | Denominator Term | Expected Mean Square |
|---|---|---|---|---|
| A (C2) | 2 | Yes | B(A) | S+csB+bcsA |
| B(A) | 3 | No | S(ABC) | S+csB |
| C (C4) | 1 | Yes | BC(A) | S+sBC+absC |
| AC | 2 | Yes | BC(A) | S+sBC+bsAC |
| BC(A) | 3 | No | S(ABC) | S+sBC |
| S(ABC) | 0 | No | | S |
| | S | | | |

Note: Expected Mean Squares are for the balanced cell-frequency case.

**Analysis of Variance Table**

| Source Term | DF | Sum of Squares | Mean Square | F-Ratio | Prob Level | Power (Alpha=0.05) |
|---|---|---|---|---|---|---|
| A (C2) | 2 | 246267.2 | 123133.6 | 5.26 | 0.104519 | 0.227807 |
| B(A) | 3 | 70225.5 | 23408.5 | | | |
| C (C4) | 1 | 37632 | 37632 | 0.31 | 0.617203 | 0.057865 |
| AC | 2 | 98376.5 | 49188.25 | 0.40 | 0.699505 | 0.063179 |
| BC(A) | 3 | 365667.5 | 121889.2 | | | |
| S | 0 | 0 | | | | |
| Total (Adjusted) | 11 | 818168.7 | | | | |
| Total | 12 | | | | | |

* Term significant at alpha = 0.05

# References

## A

**Agresti, A. and Coull, B.** 1998. “Approximate is Better than ‘Exact’ for Interval Estimation of Binomial Proportions,” *American Statistician*, Volume 52 Number 2, pages 119-126.

**A’Hern, R. P. A.** 2001. "Sample size tables for exact single-stage phase II designs." *Statistics in Medicine*, Volume 20, pages 859-866.

**AIAG (Automotive Industry Action Group)**. 1995. *Measurement Systems Analysis*. This booklet was developed by Chrysler/Ford/GM Supplier Quality Requirements Task Force. It gives a detailed discussion of how to design and analyze an R&R study. The book may be obtained from ASQC or directly from AIAG by calling 801-358-3570.

**Akaike, H.** 1973. "Information theory and an extension of the maximum likelihood principle," In B. N. Petrov & F. Csaki (Eds.), *The second international symposium on information theory*. Budapest, Hungary: Akademiai Kiado.

**Akaike, H.** 1974. "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19, (6): pages 716-723.

**Albert, A. and Harris, E**. 1987. *Multivariate Interpretation of Clinical Laboratory Data*. Marcel Dekker, New York, New York. This book is devoted to a discussion of how to apply multinomial logistic regression to medical diagnosis. It contains the algorithm that is the basis of our multinomial logistic regression routine.

**Allen, D. and Cady, F.**. 1982. *Analyzing Experimental Data by Regression*. Wadsworth. Belmont, Calif. This book works completely through several examples. It is very useful to those who want to see complete analyses of complex data.

**Al-Sunduqchi, Mahdi S.** 1990. *Determining the Appropriate Sample Size for Inferences Based on the Wilcoxon Statistics*. Ph.D. dissertation under the direction of William C. Guenther, Dept. of Statistics, University of Wyoming, Laramie, Wyoming.

**Altman, Douglas**. 1991. *Practical Statistics for Medical Research*. Chapman & Hall. New York, NY. This book provides an introductory discussion of many statistical techniques that are used in medical research. It is the only book we found that discussed ROC curves.

**Andersen, P.K., Borgan, O., Gill, R.D., and Keiding, N.** 1997. *Statistical Models Based on Counting Processess*. Springer-Verlag, New York. This is an advanced book giving many of the theoretically developments of survival analysis.

**Anderson, R.L. and Hauck, W.W.** 1983. "A new Procedure for testing equivalence in comparative bioavailability and other clinical trials." *Commun. Stat. Theory Methods.*, Volume 12, pages 2663-2692.

**Anderson, T.W. and Darling, D.A.** 1954. "A test of goodness-of-fit." *J. Amer. Statist. Assoc*, Volume 49, pages 765-769.

**Andrews, D.F., and Herzberg, A.M.** 1985. *Data*. Springer-Verlag, New York. This book is a collection of many different data sets. It gives a complete description of each.

**Armitage**. 1955. “Tests for linear trends in proportions and frequencies.” *Biometrics*, Volume 11, pages 375-386.

**Armitage, P., and Colton, T.** 1998. *Encyclopedia of Biostatistics*. John Wiley, New York.

**Armitage,P., McPherson, C.K., and Rowe, B.C.** 1969. “Repeated significance tests on accumulating data.” *Journal of the Royal Statistical Society, Series A,* 132, pages 235-244.

**Atkinson, A.C.** 1985. *Plots, Transformations, and Regression*. Oxford University Press, Oxford (also in New York). This book goes into the details of regression diagnostics and plotting. It puts together much of the recent work in this area.

**Atkinson, A.C., and Donev, A.N.** 1992. *Optimum Experimental Designs*. Oxford University Press, Oxford. This book discusses D-Optimal designs.

**Austin, P.C., Grootendorst, P., and Anderson, G.M.** 2007. “A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study,” *Statistics in Medicine*, Volume 26, pages 734-753.

## B

**Bain, L.J. and Engelhardt, M.** 1991. *Statistical Analysis of Reliability and Life-Testing Models*. Marcel Dekker. New York. This book contains details for testing data that follow the exponential and Weibull distributions.

**Baker, Frank.** 1992. *Item Response Theory*. Marcel Dekker. New York. This book contains a current overview of IRT. It goes through the details, providing both formulas and computer code. It is not light reading, but it will provide you with much of what you need if you are attempting to use this technique.

**Barnard, G.A.** 1947. "Significance tests for 2 x 2 tables." *Biometrika* 34:123-138.

**Barrentine, Larry B.** 1991. *Concepts for R&R Studies*. ASQC Press. Milwaukee, Wisconsin. This is a very good applied work book on the subject of repeatability and reproducibility studies. The ISBN is 0-87389-108-2. ASQC Press may be contacted at 800-248-1946.

**Bartholomew, D.J.** 1963. "The Sampling Distribution of an Estimate Arising in Life Testing." *Technometrics*, Volume 5 No. 3, 361-374.

**Bartlett, M.S.** 1950. "Tests of significance in factor analysis." *British Journal of Psychology (Statistical Section)*, 3, 77-85.

**Bates, D. M. and Watts, D. G.** 1981. "A relative offset orthogonality convergence criterion for nonlinear least squares," *Technometrics*, Volume 23, 179-183.

**Beal, S. L.** 1987. "Asymptotic Confidence Intervals for the Difference between Two Binomial Parameters for Use with Small Samples." *Biometrics*, Volume 43, Issue 4, 941-950.

**Belsley, Kuh, and Welsch**. 1980. *Regression Diagnostics*. John Wiley & Sons. New York. This is the book that brought regression diagnostics into the main-stream of statistics. It is a graduate level treatise on the subject.

**Benjamini, Y. and Hochberg, Y.** 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 57, No. 1, 289-300.

**Bertsekas, D.P**. 1991. *Linear Network Optimization: Algorithms and Codes*. MIT Press. Cambridge, MA.

**Blackwelder, W.C.** 1993. "Sample size and power in prospective analysis of relative risk." *Statistics in Medicine*, Volume 12, 691-698.

**Blackwelder, W.C.** 1998. "Equivalence Trials." In *Encyclopedia of Biostatistics*, John Wiley and Sons. New York. Volume 2, 1367-1372.

**Bloomfield, P**. 1976. *Fourier Analysis of Time Series*. John Wiley and Sons. New York. This provides a technical introduction to fourier analysis techniques.

**Bock, R.D., Aiken, M.** 1981. "Marginal maximum likelihood estimation of item parameters. An application of an EM algorithm. *Psychometrika*, 46, 443-459.

**Bolstad, B.M., et al.** 2003. A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias. *Bioinformatics*, 19, 185-193.

**Bonett, Douglas.** 2002. "Sample Size Requirements for Testing and Estimating Coefficient Alpha." *Journal of Educational and Behavioral Statistics*, Vol. 27, pages 335-340.

**Box, G.E.P. and Jenkins, G.M.** 1976. *Time Series Analysis - Forecasting and Control*. Holden-Day.: San Francisco, California. This is the landmark book on ARIMA time series analysis. Most of the material in chapters 6 - 9 of this manual comes from this work.

**Box, G.E.P. 1949.** "A general distribution theory for a class of likelihood criteria." *Biometrika*, 1949, **36**, 317-346.

**Box, G.E.P. 1954a.** "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variable Problems: I." *Annals of Mathematical Statistics*, **25**, 290-302.

**Box, G.E.P. 1954b.** "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variable Problems: II." *Annals of Mathematical Statistics*, **25**, 484-498.

**Box, G.E.P., Hunter, S. and Hunter.** 1978. *Statistics for Experimenters*. John Wiley & Sons, New York. This is probably the leading book in the area experimental design in industrial experiments. You definitely should acquire and study this book if you plan anything but a casual acquaintance with experimental design. The book is loaded with examples and explanations.

**Breslow, N. E.** and **Day, N. E.** 1980. *Statistical Methods in Cancer Research: Volume 1. The Analysis of Case-Control Studies*. Lyon: International Agency for Research on Cancer.

**Brown, H., and Prescott, R.** 2006. *Applied Mixed Models in Medicine.* 2nd ed. John Wiley & Sons Ltd. Chichester, West Sussex, England.

**Brush, Gary G.** 1988. *Volume 12: How to Choose the Proper Sample Size*, American Society for Quality Control, 310 West Wisconsin Ave, Milwaukee, Wisconsin, 53203. This is a small workbook for quality control workers.

**Burdick, R.K. and Larsen, G.A.** 1997. "Confidence Intervals on Measures of Variability in R&R Studies." *Journal of Quality Technology, Vol. 29, No. 3, Pages 261-273*. This article presents the formulas used to construct confidence intervals in an R&R study.

**Bury, Karl.** 1999. *Statistical Distributions in Engineering.*. Cambridge University Press. New York, NY. (www.cup.org).

## C

**Cameron, A.C. and Trivedi, P.K.** 1998. *Regression Analysis of Count Data*. Cambridge University Press. New York, NY. (www.cup.org).

**Carmines, E.G. and Zeller, R.A.** 1990. *Reliability and Validity Assessment*. Sage University Paper. 07-017. Newbury Park, CA.

**Casagrande, J. T., Pike, M.C., and Smith, P. G.** 1978. "The Power Function of the "Exact" Test for Comparing Two Binomial Distributions," *Applied Statistics*, Volume 27, No. 2, pages 176-180. This article presents the algorithm upon which our Fisher's exact test is based.

**Cattell, R.B.** 1966. "The scree test for the number of factors." *Mult. Behav. Res.* 1, 245-276.

**Cattell, R.B. and Jaspers, J.** 1967. "A general plasmode (No. 30-10-5-2) for factor analytic exercises and research." *Mult. Behav. Res. Monographs*. 67-3, 1-212.

**Chambers, J.M., Cleveland, W.S., Kleiner, B., and Tukey, P.A**. 1983. *Graphicals Methods for Data Analysis.* Duxbury Press, Boston, Mass. This wonderful little book is full of examples of ways

to analyze data graphically. It gives complete (and readable) coverage to such topics as scatter plots, probability plots, and box plots. It is strongly recommended.

**Chatfield, C.** 1984. *The Analysis of Time Series*. Chapman and Hall. New York. This book gives a very readable account of both ARMA modeling and spectral analysis. We recommend it to those who wish to get to the bottom of these methods.

**Chatterjee and Price.** 1979. *Regression Analysis by Example*. John Wiley & Sons. New York. A great hands-on book for those who learn best from examples. A newer edition is now available.

**Chen, K.W.; Chow, S.C.; and Li, G.** 1997. "A Note on Sample Size Determination for Bioequivalence Studies with Higher-Order Crossover Designs" *Journal of Pharmacokinetics and Biopharmaceutics*, Volume 25, No. 6, pages 753-765.

**Chen, T. T.** 1997. "Optimal Three-Stage Designs for Phase II Cancer Clinical Trials." *Statistics in Medicine*, Volume 16, pages 2701-2711.

**Chen, Xun.** 2002. "A quasi-exact method for the confidence intervals of the difference of two independent binomial proportions in small sample cases." *Statistics in Medicine*, Volume 21, pages 943-956.

**Chow, S.C. and Liu, J.P.** 1999. *Design and Analysis of Bioavailability and Bioequivalence Studies*. Marcel Dekker. New York.

**Chow, S.C.; Shao, J.; Wang, H.** 2003. *Sample Size Calculations in Clinical Research*. Marcel Dekker. New York.

**Chow, S.-C.; Shao, J.; Wang, H.** 2008. *Sample Size Calculations in Clinical Research, Second Edition*. Chapman & Hall/CRC. Boca Raton, Florida.

**Cochran and Cox.** 1992. *Experimental Designs. Second Edition*. John Wiley & Sons. New York. This is one of the classic books on experimental design, first published in 1957.

**Cochran, W.G. and Rubin, D.B.** 1973. “Controlling bias in observational studies,” *Sankhya, Ser. A*, Volume 35, Pages 417-446.

**Cohen, Jacob.** 1988. *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, New Jersey. This is a very nice, clearly written book. There are MANY examples. It is the largest of the sample size books. It does not deal with clinical trials.

**Cohen, Jacob.** 1990. “Things I Have Learned So Far.” *American Psychologist*, December, 1990, pages 1304-1312. This is must reading for anyone still skeptical about the need for power analysis.

**Collett, D.** 1991. *Modelling Binary Data.* Chapman & Hall, New York, New York. This book covers such topics as logistic regression, tests of proportions, matched case-control studies, and so on.

**Collett, D.** 1994. *Modelling Survival Data in Medical Research.* Chapman & Hall, New York, New York. This book covers such survival analysis topics as Cox regression and log rank tests.

**Conlon, M. and Thomas, R.** 1993. “The Power Function for Fisher’s Exact Test.” *Applied Statistics*, Volume 42, No. 1, pages 258-260. This article was used to validate the power calculations of Fisher’s Exact Test in PASS. Unfortunately, we could not use the algorithm to improve the speed because the algorithm requires equal sample sizes.

**Conover, W.J.** 1971. *Practical Nonparametric Statistics*. John Wiley & Sons, Inc. New York.

**Conover, W.J., Johnson, M.E.,** and **Johnson, M.M.** 1981. *Technometrics***, 23,** 351-361**.**

**Cook, D. and Weisberg, S.** 1982. *Residuals and Influence in Regression*. Chapman and Hall. New York. This is an advanced text in the subject of regression diagnostics.

**Cooley, W.W. and Lohnes, P.R.** 1985. *Multivariate Data Analysis*. Robert F. Krieger Publishing Co. Malabar, Florida.

**Cox, D. R.** 1972. “Regression Models and life tables.” *Journal of the Royal Statistical Society, Series B*, Volume 34, Pages 187-220. This article presents the proportional hazards regression model.

**Cox, D. R.** 1975. “Contribution to discussion of Mardia (1975a).” *Journal of the Royal Statistical Society, Series B*, Volume 37, Pages 380-381.

**Cox, D.R. and Snell, E.J.** 1981. *Applied Statistics: Principles and Examples*. Chapman & Hall. London, England.

**Cureton, E.E. and D'Agostino, R.B.** 1983. *Factor Analysis - An Applied Approach*. Lawrence Erlbaum Associates. Hillsdale, New Jersey. (This is a wonderful book for those who want to learn the details of what factor analysis does. It has both the theoretical formulas and simple worked examples to make following along very easy.)

## D

**D'Agostino, R.B., Belanger, A., D'Agostino, R.B. Jr.** 1990."A Suggestion for Using Powerful and Informative Tests of Normality.", *The American Statistician*, November 1990, Volume 44 Number 4, pages 316-321. This tutorial style article discusses D'Agostino’s tests and tells how to interpret normal probability plots.

**D'Agostino, R.B., Chase, W., Belanger, A.** 1988."The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations.", *The American Statistician*, August 1988, Volume 42 Number 3, pages 198-202.

**D'Agostino, R.B. Jr.** 2004. *Tutorials in Biostatistics*. Volume 1. John Wiley & Sons. Chichester, England.

**Dallal, G.** 1986. "An Analytic Approximation to the Distribution of Lilliefors's Test Statistic for Normality," *The American Statistician*, Volume 40, Number 4, pages 294-296.

**Daniel, C. and Wood, F.** 1980. *Fitting Equations to Data*. John Wiley & Sons. New York. This book gives several in depth examples of analyzing regression problems by computer.

**Daniel, W.** 1990. *Applied Nonparametric Statistics.* 2nd ed. PWS-KENT Publishing Company. Boston.

**Davies, Owen L.** 1971. *The Design and Analysis of Industrial Experiments*. Hafner Publishing Company, New York. This was one of the first books on experimental design and analysis. It has many examples and is highly recommended.

**Davis, J. C.** 1985. *Statistics and Data Analysis in Geology*. John Wiley. New York. (A great layman's discussion of many statistical procedures, including factor analysis.)

**Davison, A.C. and Hinkley, D.V.** 1999. *Bootstrap Methods and their Applications*. Cambridge University Press. NY, NY. This book provides and detailed account of bootstrapping.

**Davison, Mark.** 1983. *Multidimensional Scaling*. John Wiley & Sons. NY, NY. This book provides a very good, although somewhat advanced, introduction to the subject.

**DeLong, E.R., DeLong, D.M., and Clarke-Pearson, D.L.** 1988. “Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach.” *Biometrics*, 44, pages 837-845.

**DeMets, D.L. and Lan, K.K.G.** 1984. “An overview of sequential methods and their applications in clinical trials.” *Communications in Statistics, Theory and Methods*, 13, pages 2315-2338.

**DeMets, D.L. and Lan, K.K.G.** 1994. “Interim analysis: The alpha spending function approach.” *Statistics in Medicine*, 13, pages 1341-1352.

**Demidenko, E.** 2004. *Mixed Models – Theory and Applications*. John Wiley & Sons. Hoboken, New Jersey.

**Desu, M. M. and Raghavarao, D.** 1990. *Sample Size Methodology*. Academic Press. New York. (Presents many useful results for determining sample sizes.)

**DeVor, Chang, and Sutherland**. 1992. *Statistical Quality Design and Control*. Macmillan Publishing. New York. This is a comprehensive textbook of SPC including control charts, process capability, and experimental design. It has many examples. 800 pages.

**Devroye, Luc**. 1986. *Non-Uniform Random Variate Generation.* Springer-Verlag. New York. This book is currently available online at http://jeff.cs.mcgill.ca/~luc/rnbookindex.html.

**Diggle, P.J., Liang, K.Y., and Zeger, S.L.** 1994. *Analysis of Longitudinal Data*. Oxford University Press. New York, New York.

**Dillon, W. and Goldstein, M.** 1984. *Multivariate Analysis - Methods and Applications*. John Wiley. NY, NY. This book devotes a complete chapter to loglinear models. It follows Fienberg's book, providing additional discussion and examples.

**Dixon, W. J. and Tukey, J. W.** 1968. "Approximate behavior of the distribution of Winsorized t," *Technometrics*, Volume 10, pages 83-98.

**Dodson, B.** 1994. *Weibull Analysis*. ASQC Quality Press. Milwaukee, Wisconsin. This paperback book provides the basics of Weibull fitting. It contains many of the formulas used in our Weibull procedure.

**Donnelly, Thomas G.** 1980. "ACM Algorithm 462: Bivariate Normal Distribution," *Collected Algorithms from ACM*, Volume II, New York, New York.

**Donner, Allan.** 1984. "Approaches to Sample Size Estimation in the Design of Clinical Trials--A Review," *Statistics in Medicine*, Volume 3, pages 199-214. This is a well done review of the clinical trial literature. Although it is becoming out of date, it is still a good place to start.

**Donner, A. and Klar, N.** 1996. "Statistical Considerations in the Design and Analysis of Community Intervention Trials." *The Journal of Clinical Epidemiology*, Vol. 49, No. 4, 1996, pages 435-439.

**Donner, A. and Klar, N.** 2000. *Design and Analysis of Cluster Randomization Trials in Health Research.* Arnold. London.

**Draghici, S.** 2003. *Data Analysis Tools for DNA Microarrays*. Chapman & Hall/CRC. London. This is an excellent overview of most areas of Microarray analysis.

**Draper, N.R. and Smith, H.** 1966. *Applied Regression Analysis*. John Wiley & Sons. New York. This is a classic text in regression analysis. It contains both in depth theory and applications. This text is often used in graduate courses in regression analysis.

**Draper, N.R. and Smith, H.** 1981. *Applied Regression Analysis - Second Edition*. John Wiley & Sons. New York, NY. This is a classic text in regression analysis. It contains both in-depth theory and applications. It is often used in graduate courses in regression analysis.

**Dudoit, S., Shaffer, J.P., and Boldrick, J.C.** 2003. "Multiple Hypothesis Testing in Microarray Experiments," *Statistical Science*, Volume 18, No. 1, pages 71-103.

**Dudoit, S., Yang, Y.H., Callow, M.J.,** and **Speed, T.P.** 2002. "Statistical Methods for Identifying Differentially Expressed Genes in Replicated cDNA Experiments," *Statistica Sinica*, Volume 12, pages 111-139.

**du Toit, S.H.C., Steyn, A.G.W., and Stumpf, R.H.** 1986. *Graphical Exploratory Data Analysis.* Springer-Verlag. New York. This book contains examples of graphical analysis for a broad range of topics.

**Dunn, O. J.** 1964. "Multiple comparisons using rank sums," *Technometrics*, Volume 6, pages 241-252.

**Dunnett, C. W.** 1955. "A Multiple comparison procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association*, Volume 50, pages 1096-1121.

**Dunteman, G.H.** 1989. *Principal Components Analysis*. Sage University Papers, 07-069. Newbury Park, California. Telephone (805) 499-0721. This monograph costs only $7. It gives a very good introduction to PCA.

**Dupont, William.** 1988. "Power Calculations for Matched Case-Control Studies," *Biometrics*, Volume 44, pages 1157-1168.

**Dupont, William** and **Plummer, Walton D.** 1990. "Power and Sample Size Calculations--A Review and Computer Program," *Controlled Clinical Trials*, Volume 11, pages 116-128. Documents a nice public-domain program on sample size and power analysis.

**Durbin, J. and Watson, G. S.** 1950. "Testing for Serial Correlation in Least Squares Regression - I," *Biometrika*, Volume 37, pages 409-428.

**Durbin, J. and Watson, G. S.** 1951. "Testing for Serial Correlation in Least Squares Regression - II," *Biometrika*, Volume 38, pages 159-177.

**Dyke, G.V. and Patterson, H.D.** 1952. "Analysis of factorial arrangements when the data are proportions." *Biometrics*. Volume 8, pages 1-12. This is the source of the data used in the LLM tutorial.

## E

**Eckert, Joseph K.** 1990. *Property Appraisal and Assessment Administration*. International Association of Assessing Officers. 1313 East 60th Street. Chicago, IL 60637-2892. Phone: (312) 947-2044. This is a how-to manual published by the IAAO that describes how to apply many statistical procedures to real estate appraisal and tax assessment. We strongly recommend it to those using our *Assessment Model* procedure.

**Edgington, E.** 1987. *Randomization Tests*. Marcel Dekker. New York. A comprehensive discussion of randomization tests with many examples.

**Edwards, L.K.** 1993. *Applied Analysis of Variance in the Behavior Sciences*. Marcel Dekker. New York. Chapter 8 of this book is used to validate the repeated measures module of PASS.

**Efron, B. and Tibshirani, R. J.** 1993. *An Introduction to the Bootstrap*. Chapman & Hall. New York.

**Elandt-Johnson, R.C. and Johnson, N.L.** 1980. *Survival Models and Data Analysis*. John Wiley. NY, NY. This book devotes several chapters to population and clinical life-table analysis.

**Epstein, Benjamin.** 1960. "Statistical Life Test Acceptance Procedures." *Technometrics*. Volume 2.4, pages 435-446.

**Everitt, B.S. and Dunn, G.** 1992. *Applied Multivariate Data Analysis*. Oxford University Press. New York. This book provides a very good introduction to several multivariate techniques. It helps you understand how to interpret the results.

## F

**Farrington, C. P. and Manning, G.** 1990. "Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-Zero Risk Difference or Non-Unity Relative Risk." *Statistics in Medicine*, Vol. 9, pages 1447-1454. This article contains the formulas used for the Equivalence of Proportions module in PASS.

**Feldt, L.S.; Woodruff, D.J.; & Salih, F.A.** 1987. "Statistical inference for coefficient alpha." *Applied Psychological Measurement*, Vol. 11, pages 93-103.

**Feldt, L.S.; Ankenmann, R.D.** 1999. "Determining Sample Size for a Test of the Equality of Alpha Coefficients When the Number of Part-Tests is Small." *Psychological Methods*, Vol. 4(4), pages 366-377.

**Fienberg, S.** 1985. *The Analysis of Cross-Classified Categorical Data*. MIT Press. Cambridge, Massachusetts. This book provides a very good introduction to the subject. It is a must for any serious student of the subject.

**Finney, D.** 1971. *Probit Analysis*. Cambridge University Press. New York, N.Y.

**Fisher, N.I.** 1993. *Statistical Analysis of Circular Data*. Cambridge University Press. New York, New York.

**Fisher, R.A.** 1936. "The use of multiple measurements in taxonomic problems." *Annuals of Eugenics*, Volume 7, Part II, 179-188. This article is famous because in it Fisher included the 'iris data' that is always presented when discussing discriminant analysis.

**Fleiss, Joseph L.** 1981. *Statistical Methods for Rates and Proportions*. John Wiley & Sons. New York. This book provides a very good introduction to the subject.

**Fleiss, J. L., Levin, B., Paik, M.C.** 2003. *Statistical Methods for Rates and Proportions. Third Edition.* John Wiley & Sons. New York. This book provides a very good introduction to the subject.

**Fleiss, Joseph L.** 1986. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons. New York. This book provides a very good introduction to clinical trials. It may be a bit out of date now, but it is still very useful.

**Fleming, T. R.** 1982. "One-sample multiple testing procedure for Phase II clinical trials." *Biometrics*, Volume 38, pages 143-151.

**Flury, B. and Riedwyl, H.** 1988. *Multivariate Statistics: A Practical Approach*. Chapman and Hall. New York. This is a short, paperback text that provides lots of examples.

**Flury, B.** 1988. *Common Principal Components and Related Multivariate Models*. John Wiley & Sons. New York. This reference describes several advanced PCA procedures.

## G

**Gans.** 1984. "The Search for Significance: Different Tests on the Same Data." *The Journal of Statistical Computation and Simulation*, 1984, pages 1-21.

**Gart, John J. and Nam, Jun-mo.** 1988. "Approximate Interval Estimation of the Ratio in Binomial Parameters: A Review and Corrections for Skewness." *Biometrics*, Volume 44, Issue 2, 323-338.

**Gart, John J. and Nam, Jun-mo.** 1990. "Approximate Interval Estimation of the Difference in Binomial Parameters: Correction for Skewness and Extension to Multiple Tables." *Biometrics*, Volume 46, Issue 3, 637-643.

**Gehlback, Stephen.** 1988. *Interpreting the Medical Literature: Practical Epidemiology for Clinicians*. Second Edition. McGraw-Hill. New York. Telephone: (800)722-4726. The preface of this book states that its purpose is to provide the reader with a useful approach to interpreting the quantitative results given in medical literature. We reference it specifically because of its discussion of ROC curves.

**Gentle, James E.** 1998. *Random Number Generation and Monte Carlo Methods*. Springer. New York.

**Gibbons, J.** 1976. *Nonparametric Methods for Quantitative Analysis*. Holt, Rinehart and Winston. New York.

**Gleason, T.C. and Staelin, R.** 1975. "A proposal for handling missing data." *Psychometrika*, 40, 229-252.

**Goldstein, Richard.** 1989. "Power and Sample Size via MS/PC-DOS Computers," *The American Statistician*, Volume 43, Number 4, pages 253-260. A comparative review of power analysis software that was available at that time.

**Gomez, K.A. and Gomez, A. A.** 1984. *Statistical Procedures for Agricultural Research*. John Wiley & Sons. New York. This reference contains worked-out examples of many complex ANOVA designs. It includes split-plot designs. We recommend it.

**Graybill, Franklin.** 1961. *An Introduction to Linear Statistical Models*. McGraw-Hill. New York, New York. This is an older book on the theory of linear models. It contains a few worked examples of power analysis.

**Greenacre, M.** 1984. *Theory and Applications of Correspondence Analysis*. Academic Press. Orlando, Florida. This book goes through several examples. It is probably the most complete book in English on the subject.

**Greenacre, Michael J.** 1993. *Correspondence Analysis in Practice*. Academic Press. San Diego, CA. This book provides a self-teaching course in correspondence analysis. It is the clearest exposition on the subject that I have every seen. If you want to gain an understanding of CA, you must obtain this (paperback) book.

**Griffiths, P. and Hill, I.D.** 1985. *Applied Statistics Algorithms*, The Royal Statistical Society, London, England. See page 243 for ACM algorithm 291.

**Gross and Clark** 1975. *Survival Distributions*: Reliability Applications in Biomedical Sciences. John Wiley, New York.

**Gu, X.S., and Rosenbaum, P.R.** 1993. "Comparison of Multivariate Matching Methods: Structures, Distances and Algorithms," *Journal of Computational and Graphical Statistics*, Vol. 2, No. 4, pages 405-420.

**Guenther, William C.** 1977. "Desk Calculation of Probabilities for the Distribution of the Sample Correlation Coefficient," *The American Statistician*, Volume 31, Number 1, pages 45-48.

**Guenther, William C.** 1977. *Sampling Inspection in Statistical Quality Control*. Griffin's Statistical Monographs, Number 37. London.

## H

**Haberman, S.J.** 1972. "Loglinear Fit of Contingency Tables." *Applied Statistics*. Volume 21, pages 218-225. This lists the fortran program that is used to create our LLM algorithm.

**Hahn, G. J. and Meeker, W.Q.** 1991. *Statistical Intervals*. John Wiley & Sons. New York.

**Hambleton, R.K; Swaminathan, H; Rogers, H.J.** 1991. *Fundamentals of Item Response Theory*. Sage Publications. Newbury Park, California. Phone: (805)499-0721. Provides an inexpensive, readable introduction to IRT. A good place to start.

**Hamilton, L.** 1991. *Regression with Graphics: A Second Course in Applied Statistics*. Brooks/Cole Publishing Company. Pacific Grove, California. This book gives a great introduction to the use of graphical analysis with regression. It is a must for any serious user of regression. It is written at an introductory level.

**Hand, D.J. and Taylor, C.C.** 1987. *Multivariate Analysis of Variance and Repeated Measures*. Chapman and Hall. London, England.

**Hanley, J. A. and McNeil, B. J.** 1982. "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve." *Radiology*, 143, 29-36. April, 1982.

**Hanley, J. A. and McNeil, B. J.** 1983. "A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases." *Radiology*, 148, 839-843. September, 1983.

**Hartigan, J.** 1975. *Clustering Algorithms*. John Wiley. New York. (This is the "bible" of cluster algorithms. Hartigan developed the K-means algorithm used in **NCSS**.)

**Haupt, R.L. and Haupt, S.E.** 1998. *Practical Genetic Algorithms*. John Wiley. New York.

**Hernandez-Bermejo, B. and Sorribas, A.** 2001. "Analytical Quantile Solution for the S-distribution, Random Number Generation and Statistical Data Modeling." *Biometrical Journal* 43, 1007-1025.

**Hintze, J. L. and Nelson, R.D.** 1998. "Violin Plots: A Box Plot-Density Trace Synergism." *The American Statistician* 52, 181-184.

**Hoaglin, Mosteller, and Tukey.** 1985. *Exploring Data Tables, Trends, and Shapes*. John Wiley. New York.

**Hoaglin, Mosteller, and Tukey.** 1983. *Understanding Robust and Exploratory Data Analysis*. John Wiley & Sons. New York.

**Hochberg, Y. and Tamhane, A. C.** 1987. *Multiple Comparison Procedures*. John Wiley & Sons. New York.

**Hoerl, A.E. and Kennard, R.W.** 1970. "Ridge Regression: Biased estimation for nonorthogonal problems." *Technometrics* 12, 55-82.

**Hoerl, A.E. and Kennard R.W.** 1976. "Ridge regression: Iterative estimation of the biasing parameter." *Communications in Statistics* A5, 77-88.

**Howe, W.G.** 1969. "Two-Sided Tolerance Limits for Normal Populations—Some Improvements." *Journal of the American Statistical Association*, 64, 610-620.

**Hosmer, D. and Lemeshow, S.** 1989. *Applied Logistic Regression*. John Wiley & Sons. New York. This book gives an advanced, in depth look at logistic regression.

**Hosmer, D. and Lemeshow, S.** 1999. *Applied Survival Analysis*. John Wiley & Sons. New York.

**Hotelling, H.** 1933. "Analysis of a complex of statistical variables into principal components." *Journal of Educational Psychology* 24, 417-441, 498-520.

**Hsieh, F.Y.** 1989. "Sample Size Tables for Logistic Regression," *Statistics in Medicine*, Volume 8, pages 795-802. This is the article that was the basis for the sample size calculations in logistic regression in PASS 6.0. It has been superceded by the 1998 article.

**Hsieh, F.Y., Block, D.A., and Larsen, M.D.** 1998. "A Simple Method of Sample Size Calculation for Linear and Logistic Regression," *Statistics in Medicine*, Volume 17, pages 1623-1634. The sample size calculation for logistic regression in PASS are based on this article.

**Hsieh, F.Y. and Lavori, P.W.** 2000. "Sample-Size Calculations for the Cox Proportional Hazards Regression Model with Nonbinary Covariates," *Controlled Clinical Trials*, Volume 21, pages 552-560. The sample size calculation for Cox regression in PASS are based on this article.

**Hsu, Jason.** 1996. *Multiple Comparisons: Theory and Methods*. Chapman & Hall. London. This book gives a beginning to intermediate discussion of multiple comparisons, stressing the interpretation of the various MC tests. It provides details of three popular MC situations: all pairs, versus the best, and versus a control. The power calculations used in the MC module of PASS came from this book.

## I

**Irizarry, R.A., et al.** 2003a. Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data. *Biostatistics*, 4, 249-264.

**Irizarry, R.A., et al.** 2003b. Summaries of Affymetrix GeneChip Probe Level Data. *Nucleic Acids Research*, 31, e15.

## J

**Jackson, J.E.** 1991. *A User's Guide To Principal Components.* John Wiley & Sons. New York. This is a great book to learn about PCA from. It provides several examples and treats everything at a level that is easy to understand.

**James, Mike.** 1985. *Classification Algorithms*. John Wiley & Sons. New York. This is a great text on the application of discriminant analysis. It includes a simple, easy-to-understand, theoretical development as well as discussions of the application of discriminant analysis.

**Jammalamadaka, S.R. and SenGupta, A.** 2001. *Topics in Circular Statistics*. World Scientific. River Edge, New Jersey.

**Jobson, J.D.** 1992. *Applied Multivariate Data Analysis - Volume II: Categorical and Multivariate Methods*. Springer-Verlag. New York. This book is a useful reference for loglinear models and other multivariate methods. It is easy to follows and provides lots of examples.

**Jolliffe, I.T.** 1972. "Discarding variables in a principal component analysis, I: Artifical data." *Applied Statistics*, 21:160-173.

**Johnson, N.L., Kotz, S., and Kemp, A.W.** 1992. *Univariate Discrete Distributions, Second Edition*. John Wiley & Sons. New York.

**Johnson, N.L., Kotz, S., and Balakrishnan, N.** 1994. *Continuous Univariate Distributions Volume 1, Second Edition*. John Wiley & Sons. New York.

**Johnson, N.L., Kotz, S., and Balakrishnan, N.** 1995. *Continuous Univariate Distributions Volume 2, Second Edition*. John Wiley & Sons. New York.

**Jolliffe, I.T.** 1986. *Principal Component Analysis*. Springer-Verlag. New York. This book provides an easy-reading introduction to PCA. It goes through several examples.

**Julious, Steven A.** 2004. "Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data." *Statistics in Medicine*, 23:1921-1986.

**Jung, S.-H.** 2005. "Sample size for FDR-control in microarray data analysis" *Bioinformatics*, 21(14):3097-3104.

**Juran, J.M.** 1979. *Quality Control Handbook*. McGraw-Hill. New York.

## K

**Kaiser, H.F.** 1960. "The application of electronic computers to factor analysis." *Educational and Psychological Measurement*. 20:141-151.

**Kalbfleisch, J.D. and Prentice, R.L.** 1980. *The Statistical Analysis of Failure Time Data*. John Wiley, New York.

**Karian, Z.A and Dudewicz, E.J.** 2000. *Fitting Statistical Distributions.* CRC Press, New York.

**Kaufman, L. and Rousseeuw, P.J.** 1990. *Finding Groups in Data*. John Wiley. New York. This book gives an excellent introduction to cluster analysis. It treats the forming of the distance matrix and several different types of cluster methods, including fuzzy. All this is done at an elementary level so that users at all levels can gain from it.

**Kay, S.M.** 1988. *Modern Spectral Estimation.* Prentice-Hall: Englewood Cliffs, New Jersey. A very technical book on spectral theory.

**Kendall,M. and Ord, J.K.** 1990. *Time Series.* Oxford University Press. New York. This is theoretical introduction to time series analysis that is very readable.

**Kendall,M. and Stuart, A.** 1987. *Kendall's Advanced Theory of Statistics. Volume 1: Distribution Theory*. Oxford University Press. New York. This is a fine math-stat book for graduate students in statistics. We reference it because it includes formulas that are used in the program.

**Kenward, M. G. and Roger, J. H.** 1997. "Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood," *Biometrics*, 53, pages 983-997.

**Keppel, Geoffrey.** 1991. *Design and Analysis - A Researcher's Handbook.* Prentice Hall. Englewood Cliffs, New Jersey. This is a very readable primer on the topic of analysis of variance. Recommended for those who want the straight scoop with a few, well-chosen examples.

**Kirk, Roger E.** 1982. *Experimental Design: Procedures for the Behavioral Sciences.* Brooks/Cole. Pacific Grove, California. This is a respected reference on experimental design and analysis of variance.

**Klein, J.P. and Moeschberger, M.L..** 1997. *Survival Analysis.* Springer-Verlag. New York. This book provides a comprehensive look at the subject complete with formulas, examples, and lots of useful comments. It includes all the more recent developments in this field. I recommend it.

**Koch, G.G.; Atkinson, S.S.; Stokes, M.E.** 1986. *Encyclopedia of Statistical Sciences.* Volume 7. John Wiley. New York. Edited by Samuel Kotz and Norman Johnson. The article on Poisson Regression provides a very good summary of the subject.

**Kotz and Johnson.** 1993. *Process Capability Indices.* Chapman & Hall. New York. This book gives a detailed account of the capability indices used in SPC work. 207 pages.

**Kraemer, H. C.** and **Thiemann, S.** 1987. *How Many Subjects*, Sage Publications, 2111 West Hillcrest Drive, Newbury Park, CA. 91320. This is an excellent introduction to power analysis.

**Kruskal, J.** 1964. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis." *Psychometrika* 29, pages 1-27, 115-129. This article presents the algorithm on which the non-metric algorithm used in NCSS is based.

**Kruskal, J. and Wish, M.** 1978. *Multidimensional Scaling*. Sage Publications. Beverly Hills, CA. This is a well-written monograph by two of the early pioneers of MDS. We suggest it to all serious students of MDS.

**Kuehl, R.O.** 2000. *Design of Experiment: Statistical Principles of Research Design and Analysis, 2nd Edition.* Brooks/Cole. Pacific Grove, California. This is a good graduate level text on experimental design with many examples.

## L

**Lachenbruch, P.A.** 1975. *Discriminant Analysis.* Hafner Press. New York. This is an in-depth treatment of the subject. It covers a lot of territory, but has few examples.

**Lachin, John M.** 2000. *Biostatistical Methods.* John Wiley & Sons. New York. This is a graduate-level methods book that deals with statistical methods that are of interest to biostatisticians such as odds ratios, relative risks, regression analysis, case-control studies, and so on.

**Lachin, John M.** and **Foulkes, Mary A. 1986.** "Evaluation of Sample Size and Power for Analyses of Survival with Allowance for Nonuniform Patient Entry, Losses to Follow-up, Noncompliance, and Stratification," *Biometrics*, Volume 42, September, pages 507-516.

**Lan, K.K.G. and DeMets, D.L.** 1983. "Discrete sequential boundaries for clinical trials." *Biometrika*, 70, pages 659-663.

**Lan, K.K.G. and Zucker, D.M.** 1993. "Sequential monitoring of clinical trials: the role of information and Brownian motion." *Statistics in Medicine*, 12, pages 753-765.

**Lance, G.N. and Williams, W.T.** 1967. "A general theory of classificatory sorting strategies. I. Hierarchical systems." *Comput. J.* 9, pages 373-380.

**Lance, G.N. and Williams, W.T.** 1967. "Mixed-data classificatory programs I. Agglomerative systems." *Aust.Comput. J.* 1, pages 15-20.

**Lawless, J.F.** 1982. *Statistical Models and Methods for Lifetime Data*. John Wiley, New York.

**Lawson, John.** 1987. *Basic Industrial Experimental Design Strategies*. Center for Statistical Research at Brigham Young University. Provo, Utah. 84602. This is a manuscript used by Dr. Lawson in courses and workshops that he provides to industrial engineers. It is the basis for many of our experimental design procedures.

**Lebart, Morineau, and Warwick.** 1984. *Multivariate Descriptive Statistical Analysis*. John Wiley & Sons. This book devotes a large percentage of its discussion to correspondence analysis.

**Lee, E.T.** 1974. "A Computer Program for Linear Logistic Regression Analysis" in *Computer Programs in Biomedicine*, Volume 4, pages 80-92.

**Lee, E.T.** 1980. *Statistical Methods for Survival Data Analysis*. Lifetime Learning Publications. Belmont, California.

**Lee, E.T.** 1992. *Statistical Methods for Survival Data Analysis*. Second Edition. John Wiley & Sons. New York. This book provides a very readable introduction to survival analysis techniques.

**Lee, M.-L. T.** 2004. *Analysis of Microarray Gene Expression Data*. Kluwer Academic Publishers. Norwell, Massachusetts.

**Lee, S. K.** 1977. "On the Asymptotic Variances of u Terms in Loglinear Models of Multidimensional Contingency Tables." *Journal of the American Statistical Association*. Volume 72 (June, 1977), page 412. This article describes methods for computing standard errors that are used in the LLM section of this program.

**Lenth, Russell V.** 1987. "Algorithm AS 226: Computing Noncentral Beta Probabilities," *Applied Statistics*, Volume 36, pages 241-244.

**Lenth, Russell V.** 1989. "Algorithm AS 243: Cumulative Distribution Function of the Non-central t Distribution," *Applied Statistics*, Volume 38, pages 185-189.

**Lesaffre, E. and Albert, A.** 1989. "Multiple-group Logistic Regression Diagnostics" *Applied Statistics*, Volume 38, pages 425-440. See also Pregibon 1981.

**Levene, H.** 1960. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, I. Olkin et al., eds. Stanford University Press, Stanford Calif., pp. 278-292.

**Lewis, J.A.** 1999. "Statistical principles for clinical trials (ICH E9) an introductory note on an international guideline." *Statistics in Medicine*, 18, pages 1903-1942.

**Lipsey, Mark W.** 1990. *Design Sensitivity Statistical Power for Experimental Research*, Sage Publications, 2111 West Hillcrest Drive, Newbury Park, CA. 91320. This is an excellent introduction to power analysis.

**Little, R. and Rubin, D.** 1987. *Statistical Analysis with Missing Data*. John Wiley & Sons. New York. This book is completely devoted to dealing with missing values. It gives a complete treatment of using the EM algorithm to estimate the covariance matrix.

**Little, R. C. et al.** 2006. *SAS for Mixed Models – Second Edition*. SAS Institute Inc., Cary, North Carolina.

**Liu, H. and Wu, T. 2005.** "Sample Size Calculation and Power Analysis of Time-Averaged Difference," *Journal of Modern Applied Statistical Methods*, Vol. 4, No. 2, pages 434-445.

**Lu, Y. and Bean, J.A.** 1995. "On the sample size for one-sided equivalence of sensitivities based upon McNemar's test," *Statistics in Medicine*, Volume 14, pages 1831-1839.

**Lui, J., Hsueh, H., Hsieh, E., and Chen, J.J.** 2002. "Tests for equivalence or non-inferiority for paired binary data," *Statistics in Medicine*, Volume 21, pages 231-245.

**Lloyd, D.K. and Lipow, M.** 1991. *Reliability: Management, Methods, and Mathematics*. ASQC Quality Press. Milwaukee, Wisconsin.

**Locke, C.S.** 1984. "An exact confidence interval for untransformed data for the ratio of two formulation means," *J. Pharmacokinet. Biopharm.*, Volume 12, pages 649-655.

**Lockhart, R. A. & Stephens, M. A.** 1985. "Tests of fit for the von Mises distribution." *Biometrika* 72, pages 647-652.

# M

**Machin, D., Campbell, M., Fayers, P., and Pinol, A.** 1997. *Sample Size Tables for Clinical Studies, 2nd Edition*. Blackwell Science. Malden, Mass. A very good & easy to read book on determining appropriate sample sizes in many situations.

**Makridakis, S. and Wheelwright, S.C.** 1978. *Iterative Forecasting*. Holden-Day.: San Francisco, California. This is a very good book for the layman since it includes several detailed examples. It is written for a person with a minimum amount of mathematical background.

**Manly, B.F.J.** 1986. *Multivariate Statistical Methods - A Primer*. Chapman and Hall. New York. This nice little paperback provides a simplified introduction to many multivariate techniques, including MDS.

**Mardia, K.V. and Jupp, P.E.** 2000. *Directional Statistics*. John Wiley & Sons. New York.

**Marple, S.L.** 1987. *Digital Spectral Analysis with Applications.* Prentice-Hall: Englewood Cliffs, New Jersey. A technical book about spectral analysis.

**Martinez and Iglewicz.** 1981. "A test for departure from normality based on a biweight estimator of scale." *Biometrika*, 68, 331-333).

**Marubini, E.** and **Valsecchi, M.G.** 1996. *Analysing Survival Data from Clinical Trials and Observational Studies*. John Wiley: New York, New York.

**Mather, Paul.** 1976. *Computational Methods of Multivariate Analysis in Physical Geography.* John Wiley & Sons. This is a great book for getting the details on several multivariate procedures. It was written for non-statisticians. It is especially useful in its presentation of cluster analysis. Unfortunately, it is out-of-print. You will have to look for it in a university library (it is worth the hunt).

**Matsumoto, M. and Nishimura,T.** 1998. "Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator" *ACM Trans. On Modeling and Computer Simulations*.

**Mauchly, J.W.** 1940. "Significance test for sphericity of a normal n-variate distribution." *Annals of Mathematical Statistics*, 11: 204-209

**McCabe, G.P.** 1984. "Principal variables." *Technometrics*, 26, 137-144.

**McClish, D.K.** 1989. "Analyzing a Portion of the ROC Curve." *Medical Decision Making*, 9: 190-195

**McHenry, Claude.** 1978. "Multivariate subset selection." *Journal of the Royal Statistical Society, Series C*. Volume 27, No. 23, pages 291-296.

**McNeil, D.R.** 1977. *Interactive Data Analysis*. John Wiley & Sons. New York.

**Mendenhall, W.** 1968. *Introduction to Linear Models and the Design and Analysis of Experiments*. Wadsworth. Belmont, Calif.

**Metz, C.E.** 1978. "Basic principles of ROC analysis." *Seminars in Nuclear Medicine*, Volume 8, No. 4, pages 283-298.

**Miettinen, O.S. and Nurminen, M.** 1985. "Comparative analysis of two rates." *Statistics in Medicine* 4: 213-226.

**Milliken, G.A. and Johnson, D.E.** 1984. *Analysis of Messy Data, Volume 1*. Van Nostrand Rienhold. New York, NY.

**Milne, P.** 1987. *Computer Graphics for Surveying.* E. & F. N. Spon, 29 West 35th St., NY, NY 10001

**Montgomery, Douglas.** 1984. *Design and Analysis of Experiments*. John Wiley & Sons, New York. A textbook covering a broad range of experimental design methods. The book is not limited to industrial investigations, but gives a much more general overview of experimental design methodology.

**Montgomery, Douglas and Peck.** 1992. *Introduction to Linear Regression Analysis*. A very good book on this topic.

**Montgomery, Douglas C.** 1991. *Introduction to Statistical Quality Control.* Second edition. John Wiley & Sons. New York. This is a comprehensive textbook of SPC including control charts, process capability, and experimental design. It has many examples. 700 pages.

**Moore, D. S. and McCabe, G. P.** 1999. *Introduction to the Practice of Statistics*. W. H. Freeman and Company. New York.

**Mosteller, F. and Tukey, J.W.** 1977. *Data Analysis and Regression*. Addison-Wesley. Menlo Park, California. This book should be read by all serious users of regression analysis. Although the terminology is a little different, this book will give you a fresh look at the whole subject.

**Motulsky, Harvey.** 1995. *Intuitive Biostatistics.* Oxford University Press. New York, New York. This is a wonderful book for those who want to understand the basic concepts of statistical testing. The author presents a very readable coverage of the most popular biostatistics tests. If you have forgotten how to interpret the various statistical tests, get this book!

**Moura, Eduardo C.** 1991. *How To Determine Sample Size And Estimate Failure Rate in Life Testing*. ASQC Quality Press. Milwaukee, Wisconsin.

**Mueller, K. E., and Barton, C. N.** 1989. "Approximate Power for Repeated-Measures ANOVA Lacking Sphericity." *Journal of the American Statistical Association,* Volume 84, No. 406, pages 549-555.

**Mueller, K. E., LaVange, L.E., Ramey, S.L., and Ramey, C.T.** 1992. "Power Calculations for General Linear Multivariate Models Including Repeated Measures Applications." *Journal of the American Statistical Association,* Volume 87, No. 420, pages 1209-1226.

**Mukerjee, H., Robertson, T., and Wright, F.T.** 1987. "Comparison of Several Treatments With a Control Using Multiple Contrasts." *Journal of the American Statistical Association,* Volume 82, No. 399, pages 902-910.

**Muller, K. E. and Stewart, P.W.** 2006. *Linear Model Theory: Univariate, Multivariate, and Mixed Models*. John Wiley & Sons Inc. Hoboken, New Jersey.

**Myers, R.H.** 1990. *Classical and Modern Regression with Applications*. PWS-Kent Publishing Company. Boston, Massachusetts. This is one of the bibles on the topic of regression analysis.

## N

**Naef, F. et al.** 2002. "Empirical characterization of the expression ratio noise structure in high-density oligonucleotide arrays," *Genome Biol.*, 3, RESEARCH0018.

**Nam, Jun-mo.** 1992. "Sample Size Determination for Case-Control Studies and the Comparison of Stratified and Unstratified Analyses," *Biometrics*, Volume 48, pages 389-395.

**Nam, Jun-mo.** 1997. "Establishing equivalence of two treatments and sample size requirements in matched-pairs design," *Biometrics*, Volume 53, pages 1422-1430.

**Nam, J-m. and Blackwelder, W.C.** 2002. "Analysis of the ratio of marginal probabilities in a matched-pair setting," *Statistics in Medicine*, Volume 21, pages 689-699.

**Nash, J. C.** 1987. *Nonlinear Parameter Estimation*. Marcel Dekker, Inc. New York, NY.

**Nash, J.C.** 1979. *Compact Numerical Methods for Computers*. John Wiley & Sons. New York, NY.

**Nel, D.G. and van der Merwe, C.A.** 1986. "A solution to the multivariate Behrens-Fisher problem." *Communications in Statistics—Series A, Theory and Methods*, 15, pages 3719-3735.

**Nelson, W.B.** 1982. *Applied Life Data Analysis*. John Wiley, New York.

**Nelson, W.B.** 1990. *Accelerated Testing*. John Wiley, New York.

**Neter, J., Kutner, M., Nachtsheim, C., and Wasserman, W.** 1996. *Applied Linear Statistical Models*. Richard D. Irwin, Inc. Chicago, Illinois. This mammoth book covers regression analysis and analysis of variance thoroughly and in great detail. We recommend it.

**Neter, J., Wasserman, W., and Kutner, M**. 1983. *Applied Linear Regression Models*. Richard D. Irwin, Inc. Chicago, Illinois. This book provides you with a complete introduction to the methods of regression analysis. We suggest it to non-statisticians as a great reference tool.

**Newcombe, Robert G.** 1998a. "Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods." *Statistics in Medicine*, Volume 17, 857-872.

**Newcombe, Robert G.** 1998b. "Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods." *Statistics in Medicine*, Volume 17, 873-890.

**Newcombe, Robert G.** 1998c. "Improved Confidence Intervals for the Difference Between Binomial Proportions Based on Paired Data." *Statistics in Medicine*, Volume 17, 2635-2650.

**Newton, H.J.** 1988. *TIMESLAB: A Time Series Analysis Laboratory*. Wadsworth & Brooks/Cole: Pacific Grove, California. This book is loaded with theoretical information about time series analysis. It includes software designed by Dr. Newton for performing advanced time series and spectral analysis. The book requires a strong math and statistical background.

## O

**O'Brien, P.C. and Fleming, T.R.** 1979. "A multiple testing procedure for clinical trials." *Biometrics*, 35, pages 549-556.

**O'Brien, R.G. and Kaiser, M.K.** 1985. "MANOVA Method for Analyzing Repeated Measures Designs: An Extensive Primer." *Psychological Bulletin*, 97, pages 316-333.

**Obuchowski, N.** 1998. "Sample Size Calculations in Studies of Test Accuracy." *Statistical Methods in Medical Research*, 7, pages 371-392.

**Obuchowski, N. and McClish, D.** 1997. "Sample Size Determination for Diagnostic Accuracy Studies Involving Binormal ROC Curve Indices." *Statistics in Medicine*, 16, pages 1529-1542.

**Odeh, R.E. and Fox, M.** 1991. *Sample Size Choice*. Marcel Dekker, Inc. New York, NY.

**O'Neill and Wetherill.** 1971 "The Present State of Multiple Comparison Methods," *The Journal of the Royal Statistical Society*, Series B, vol.33, 218-250).

**Orloci, L. & Kenkel, N.** 1985. *Introduction to Data Analysis*. International Co-operative Publishing House. Fairland, Maryland. This book was written for ecologists. It contains samples and BASIC programs of many statistical procedures. It has one brief chapter on MDS, and it includes a non-metric MDS algorithm.

**Ostle, B.** 1988. *Statistics in Research. Fourth Edition*. Iowa State Press. Ames, Iowa. A comprehension book on statistical methods.

**Ott, L.** 1977. *An Introduction to Statistical Methods and Data Analysis*. Wadsworth. Belmont, Calif. Use the second edition.

**Ott, L.** 1984. *An Introduction to Statistical Methods and Data Analysis, Second Edition*. Wadsworth. Belmont, Calif. This is a complete methods text. Regression analysis is the focus of five or six chapters. It stresses the interpretation of the statistics rather than the calculation, hence it provides a good companion to a statistical program like ours.

**Owen, Donald B.** 1956. "Tables for Computing Bivariate Normal Probabilities," *Annals of Mathematical Statistics*, Volume 27, pages 1075-1090.

**Owen, Donald B.** 1965. "A Special Case of a Bivariate Non-Central t-Distribution," *Biometrika*, Volume 52, pages 437-446.

## P

**Pandit, S.M. and Wu, S.M.** 1983. *Time Series and System Analysis with Applications*. John Wiley and Sons. New York. This book provides an alternative to the Box-Jenkins approach for dealing with ARMA models. We used this approach in developing our automatic ARMA module.

**Parmar, M.K.B. and Machin, D.** 1995. *Survival Analysis*. John Wiley and Sons. New York.

**Parmar, M.K.B., Torri, V., and Steart, L.** 1998. "Extracting Summary Statistics to Perform Meta-Analyses of the Published Literature for Survival Endpoints." *Statistics in Medicine* 17, 2815-2834.

**Pearson, K.** 1901. "On lines and planes of closest fit to a system of points in space." *Philosophical Magazine* 2, 557-572.

**Pan, Z. and Kupper, L.** 1999. "Sample Size Determination for Multiple Comparison Studies Treating Confidence Interval Width as Random." *Statistics in Medicine* 18, 1475-1488.

**Pedhazur, E.L. and Schmelkin, L.P.** 1991. *Measurement, Design, and Analysis: An Integrated Approach.* Lawrence Erlbaum Associates. Hillsdale, New Jersey. This mammoth book (over 800 pages) covers multivariate analysis, regression analysis, experimental design, analysis of variance, and much more. It provides annotated output from SPSS and SAS which is also useful to our users. The text is emphasizes the social sciences. It provides a "how-to," rather than a theoretical, discussion. Its chapters on factor analysis are especially informative.

**Phillips, Kem F.** 1990. "Power of the Two One-Sided Tests Procedure in Bioequivalence," *Journal of Pharmacokinetics and Biopharmaceutics*, Volume 18, No. 2, pages 137-144.

**Pocock, S.J.** 1977. "Group sequential methods in the design and analysis of clinical trials." *Biometrika*, 64, pages 191-199.

**Press, S. J. and Wilson, S.** 1978. "Choosing Between Logistic Regression and Discriminant Analysis." *Journal of the American Statistical Association*, Volume 73, Number 364, Pages 699-705. This article details the reasons why logistic regression should be the preferred technique.

**Press, William H.** 1986. *Numerical Recipes*, Cambridge University Press, New York, New York.

**Pregibon, Daryl.** 1981. "Logistic Regression Diagnostics." *Annals of Statistics*, Volume 9, Pages 705-725. This article details the extensions of the usual regression diagnostics to the case of logistic regression. These results were extended to multiple-group logistic regression in Lesaffre and Albert (1989).

**Price, K., Storn R., and Lampinen, J.** 2005. *Differential Evolution – A Practical Approach to Global Optimization.* Springer. Berlin, Germany.

**Prihoda, Tom.** 1983. "Convenient Power Analysis For Complex Analysis of Variance Models." *Poster Session of the American Statistical Association Joint Statistical Meetings*, August 15-18, 1983, Toronto, Canada. Tom is currently at the University of Texas Health Science Center. This article includes FORTRAN code for performing power analysis.

## R

**Ramsey, Philip H.** 1978 "Power Differences Between Pairwise Multiple Comparisons," *JASA*, vol. 73, no. 363, pages 479-485.

**Rao, C.R. , Mitra, S.K., & Matthai, A.** 1966. *Formulae and Tables for Statistical Work.* Statistical Publishing Society, Indian Statistical Institute, Calcutta, India.

**Ratkowsky, David A.** 1989. *Handbook of Nonlinear Regression Models*. Marcel Dekker. New York. A good, but technical, discussion of various nonlinear regression models.

**Rawlings John O.** 1988. *Applied Regression Analysis: A Research Tool*. Wadsworth. Belmont, California. This is a readable book on regression analysis. It provides a thorough discourse on the subject.

**Reboussin, D.M., DeMets, D.L., Kim, K, and Lan, K.K.G.** 1992. "Programs for computing group sequential boundaries using the Lan-DeMets Method." Technical Report 60, Department of Biostatistics, University of Wisconsin-Madison.

**Rencher, Alvin C.** 1998. *Multivariate Statistical Inference and Applications*. John Wiley. New York, New York. This book provides a comprehensive mixture of theoretical and applied results in multivariate analysis. My evaluation may be biased since Al Rencher took me fishing when I was his student.

**Robins, Greenland, and Breslow.** 1986. "A General Estimator for the Variance of the Mantel-Haenszel Odds Ratio," *American Journal of Epidemiology*, vol.42, pages 719-723.

**Robins, Breslow, and Greenland.** 1986. "Estimators of the Mantel-Haenszel variance consisten in both sparse data and large-strata limiting models," *Biometrics*, vol. 42, pages 311-323.

**Rosenbaum, P.R.** 1989. "Optimal Matching for Observational Studies," *Journal of the American Statistical Association*, vol. 84, no. 408, pages 1024-1032.

**Rosenbaum, P.R., and Rubin, D.B.** 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, pages 41-55.

**Rosenbaum, P.R., and Rubin, D.B.** 1984. "Reducing bias in observational studies using subclassification on the propensity score," *Journal of the American Statistical Association*, vol. 79, pages 516-524.

**Rosenbaum, P.R., and Rubin, D.B.** 1985a. "Constructing a Control Group using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *American Statistician*, vol. 39, pages 33-38.

**Rosenbaum, P.R., and Rubin, D.B.** 1985b. "The Bias Due to Incomplete Matching," *Biometrics*, vol. 41, pages 106-116.

**Ryan, Thomas P.** 1989. *Statistical Methods for Quality Improvement*. John Wiley & Sons. New York. This is a comprehensive treatment of SPC including control charts, process capability, and experimental design. It provides many rules-of-thumb and discusses many non-standard situations. This is a very good 'operators manual' type of book. 446 pages.

**Ryan, Thomas P.** 1997. *Modern Regression Methods*. John Wiley & Sons. New York. This is a comprehensive treatment of regression analysis. The author often deals with practical issues that are left out of other texts.

## S

**Sahai, Hardeo & Khurshid, Anwer.** 1995. *Statistics in Epidemiology*. CRC Press. Boca Raton, Florida.

**Schiffman, Reynolds, & Young.** 1981. *Introduction to Multidimensional Scaling*. Academic Press. Orlando, Florida. This book goes through several examples.

**Schilling, Edward.** 1982. *Acceptance Sampling in Quality Control*. Marcel-Dekker. New York.

**Schlesselman, Jim.** 1981. *Case-Control Studies*. Oxford University Press. New York. This presents a complete overview of case-control studies. It was our primary source for the Mantel-Haenszel test.

**Schmee and Hahn.** November, 1979. "A Simple Method for Regression Analysis." *Technometrics*, Volume 21, Number 4, pages 417-432.

**Schoenfeld, David A.** 1983. "Sample-Size Formula for the Proportional-Hazards Regression Model" *Biometrics,* Volume 39, pages 499-503.

**Schoenfeld, David A.** and **Richter, Jane R.** 1982**.** "Nomograms for Calculating the Number of Patients Needed for a Clinical Trial with Survival as an Endpoint," *Biometrics,* March 1982, Volume 38, pages 163-170.

**Schork, M. and Williams, G.** 1980. "Number of Observations Required for the Comparison of Two Correlated Proportions." *Communications in Statistics-Simula. Computa.,* B9(4), 349-357.

**Schuirmann, Donald.** 1981**.** "On hypothesis testing to determine if the mean of a normal distribution is continued in a known interval," *Biometrics,* Volume 37, pages 617.

**Schuirmann, Donald.** 1987**.** "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability," *Journal of Pharmacokinetics and Biopharmaceutics,* Volume 15, Number 6, pages 657-680.

**Seber, G.A.F.** 1984. *Multivariate Observations*. John Wiley & Sons. New York. (This book is an encyclopedia of multivariate techniques. It emphasizes the mathematical details of each technique and provides a complete set of references. It will only be useful to those comfortable with reading mathematical equations based on matrices.)

**Seber, G.A.F. and Wild, C.J.** 1989. *Nonlinear Regression*. John Wiley & Sons. New York. This book is an encyclopedia of nonlinear regression.

**Senn, Stephen.** 1993. *Cross-over Trials in Clinical Research*. John Wiley & Sons. New York.

**Senn, Stephen.** 2002. *Cross-over Trials in Clinical Research.* Second Edition. John Wiley & Sons. New York.

**Shapiro, S.S. and Wilk, M.B.** 1965 "An analysis of Variance test for normality." *Biometrika*, Volume 52, pages 591-611.

**Shuster, Jonathan J.** 1990. *CRC Handbook of Sample Size Guidelines for Clinical Trials.* CRC Press, Boca Raton, Florida. This is an expensive book ($300) of tables for running log-rank tests. It is well documented, but at this price it better be.

**Signorini, David.** 1991. "Sample size for Poisson regression," *Biometrika,* Volume 78, 2, pages 446-450.

**Simon, Richard.** "Optimal Two-Stage Designs for Phase II Clinical Trials," *Controlled Clinical Trials,* 1989, Volume 10, pages 1-10.

**Snedecor, G. and Cochran, Wm.** 1972. *Statistical Methods*. The Iowa State University Press. Ames, Iowa.

**Sorribas, A., March, J., and Trujillano, J.** 2002. "A new parametric method based on S-distributions for computing receiver operating characteristic curves for continuous diagnostic tests." *Statistics in Medicine* 21, 1213-1235.

**Spath, H.** 1985. *Cluster Dissection and Analysis.* Halsted Press. New York. (This book contains a detailed discussion of clustering techniques for large data sets. It contains some heavy mathematical notation.)

**Speed, T.P. (editor).** 2003. *Statistical Analysis of Gene Expression Microarray Data.* Chapman & Hall/CRC. Boca Raton, Florida.

**Stekel, D.** 2003. *Microarray Bioinformatics.* Cambridge University Press. Cambridge, United Kingdom.

**Sutton, A.J., Abrams, K.R., Jones, D.R., Sheldon, T.A., and Song, F.** 2000. *Methods for Meta-Analysis in Medical Research.* John Wiley & Sons. New York.

**Swets, John A.** 1996. *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics - Collected Papers.* Lawrence Erlbaum Associates. Mahway, New Jersey.

## T

**Tabachnick, B. and Fidell, L.** 1989. *Using Multivariate Statistics.* Harper Collins. 10 East 53d Street, NY, NY 10022. This is an extremely useful text on multivariate techniques. It presents computer printouts and discussion from several popular programs. It provides checklists for each procedure as well as sample written reports. I strongly encourage you to obtain this book!

**Tango, Toshiro.** 1998. "Equivalence Test and Confidence Interval for the Difference in Proportions for the Paired-Sample Design." *Statistics in Medicine*, Volume 17, 891-908.

**Therneau, T.M. and Grambsch, P.M.** 2000. *Modeling Survival Data*. Springer: New York, New York. A the time of the writing of the Cox regression procedure, this book provides a thorough, up-to-date discussion of this procedure as well as many extensions to it. Recommended, especially to those with at least a masters in statistics.

**Thomopoulos, N.T.** 1980. *Applied Forecasting Methods*. Prentice-Hall: Englewood Cliffs, New Jersey. This book contains a very good presentation of the classical forecasting methods discussed in chapter two.

**Thompson, Simon G.** 1998. *Encyclopedia of Biostatistics, Volume 4*. John Wiley & Sons. New York. Article on Meta-Analysis on pages 2570-2579.

**Tiku, M. L.** 1965. "Laguerre Series Forms of Non-Central $X^2$ and F Distributions," *Biometrika*, Volume 42, pages 415-427.

**Torgenson, W.S.** 1952. "Multidimensional scaling. I. Theory and method." *Psychometrika* 17, 401-419. This is one of the first articles on MDS. There have been many advances, but this article presents many insights into the application of the technique. It describes the algorithm on which the metric solution used in this program is based.

**Tubert-Bitter, P., Manfredi,R., Lellouch, J., Begaud, B.** 2000. "Sample size calculations for risk equivalence testing in pharmacoepidemiology." *Journal of Clinical Epidemiology* 53, 1268-1274.

**Tukey, J.W. and McLaughlin, D.H.** 1963. "Less Vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization." *Sankhya, Series A* 25, 331-352.

**Tukey, J.W.** 1977. *Exploratory Data Analysis.* Addison-Wesley Publishing Company. Reading, Mass.

## U

**Upton, G.J.G.** 1982."A Comparison of Alternative Tests for the 2 x 2 Comparative Trial.", *Journal of the Royal Statistical Society*, Series A,, Volume 145, pages 86-105.

**Upton, G.J.G. and Fingleton, B.** 1989. *Spatial Data Analysis by Example: Categorical and Directional Data. Volume 2.* John Wiley & Sons. New York.

## V

**Velicer, W.F.** 1976. "Determining the number of components from the matrix of partial correlations." *Psychometrika*, 41, 321-327.

**Velleman, Hoaglin.** 1981. *ABC's of Exploratory Data Analysis.* Duxbury Press, Boston, Massachusetts.

**Voit, E.O.** 1992. "The S-distribution. A tool for approximation and classification of univariate, unimodal probability distributions." *Biometrical J.* 34, 855-878.

**Voit, E.O.** 2000. "A Maximum Likelihood Estimator for Shape Parameters of S-Distributions." *Biometrical J.* 42, 471-479.

**Voit, E.O. and Schwacke, L.** 1998. "Scalability properties of the S-distribution." *Biometrical J.* 40, 665-684.

**Voit, E.O. and Yu, S.** 1994. "The S-distribution. Approximation of discrete distributions." *Biometrical J.* 36, 205-219.

## W

**Walter, S.D., Eliasziw, M., and Donner, A.** 1998. "Sample Size and Optimal Designs For Reliability Studies." *Statistics in Medicine*, 17, 101-110.

**Welch, B.L.** 1938. "The significance of the difference between two means when the population variances are unequal." *Biometrika*, 29, 350-362.

**Welch, B.L.** 1947. "The Generalization of "Student's" Problem When Several Different Population Variances Are Involved," *Biometrika*, 34, 28-35.

**Welch, B.L.** 1949. "Further Note on Mrs. Aspin's Tables and on Certain Approximations to the Tabled Function," *Biometrika*, 36, 293-296.

**Westfall, P. et al.** 1999. *Multiple Comparisons and Multiple Tests Using the SAS System*. SAS Institute Inc. Cary, North Carolina.

**Westgard, J.O.** 1981. "A Multi-Rule Shewhart Chart for Quality Control in Clinical Chemistry," *Clinical Chemistry*, Volume 27, No. 3, pages 493-501. (This paper is available online at the www.westgard.com).

**Westlake, W.J.** 1981. "Bioequivalence testing—a need to rethink," *Biometrics*, Volume 37, pages 591-593.

**Whittemore, Alice.** 1981. "Sample Size for Logistic Regression with Small Response Probability," *Journal of the American Statistical Association*, Volume 76, pages 27-32.

**Wickens, T.D.** 1989. *Multiway Contingency Tables Analysis for the Social Sciences.* Lawrence Erlbaum Associates. Hillsdale, New Jersey. A thorough book on the subject. Discusses loglinear models in depth.

**Wilson, E.B..** 1927. "Probable Inference, the Law of Succession, and Statistical Inference," *Journal of the American Statistical Association*, Volume 22, pages 209-212. This article discusses the 'score' method that has become popular when dealing with proportions.

**Winer, B.J.** 1991**.** *Statistical Principles in Experimental Design (Third Edition)*. McGraw-Hill. New York, NY. A very complete analysis of variance book.

**Wit, E., and McClure, J.** 2004. *Statistics for Microarrays*. John Wiley & Sons Ltd, Chichester, West Sussex, England.

**Wolfinger, R., Tobias, R. and Sall, J.** 1994. "Computing Gaussian likelihoods and their derivatives for general linear mixed models," *SIAM Journal of Scientific Computing*, 15, no.6, pages 1294-1310.

**Woolson, R.F., Bean, J.A., and Rojas, P.B.** 1986. "Sample Size for Case-Control Studies Using Cochran's Statistic," *Biometrics*, Volume 42, pages 927-932.

## Y

**Yuen, K.K. and Dixon, W. J.** 1973. "The approximate behavior and performance of the two-sample trimmed t," *Biometrika*, Volume 60, pages 369-374.

**Yuen, K.K.** 1974. "The two-sample trimmed t for unequal population variances," *Biometrika*, Volume 61, pages 165-170.

## Z

**Zar, Jerrold H.** 1984**.** *Biostatistical Analysis (Second Edition)*. Prentice-Hall. Englewood Cliffs, New Jersey. This introductory book presents a nice blend of theory, methods, and examples for a long list of topics of interest in biostatistical work.

**Zhou, X., Obuchowski, N., McClish, D.** 2002**.** *Statistical Methods in Diagnostic Medicine*. John Wiley & Sons, Inc. New York, New York. This is a great book on the designing and analyzing diagnostic tests. It is especially useful for its presentation of ROC curves.

# Chapter Index

## 3



## A



## B



## C



## D

## E



## F



## G



## H



## I

## K



## L



## M



## N



## O



## P

## Q



## R



## S

## T



## U



## V



## W



## X

# Index

## 2



## 3



## A

## B

## C

## D

# E



# F

## G

## H



## I

## J



## K



## L

# M

## N

## O

# P

## Q



## R

# S

## T

## U

## V



## W

## X



## Y



## Z