

Chapter 605

Multivariate Analysis of Variance (MANOVA)

Introduction

This module calculates power for multivariate analysis of variance (MANOVA) designs having up to three factors. It computes power for three MANOVA test statistics: Wilks' lambda, Pillai-Bartlett trace, and Hotelling-Lawley trace.

MANOVA is an extension of common analysis of variance (ANOVA). In ANOVA, differences among various group means on a single-response variable are studied. In MANOVA, the number of response variables is increased to two or more. The hypothesis concerns a comparison of vectors of group means. The multivariate extension of the F -test is not completely direct. Instead, several test statistics are available. The actual distributions of these statistics are difficult to calculate, so we rely on approximations based on the F -distribution.

Assumptions

The following assumptions are made when using MANOVA to analyze a factorial experimental design.

1. The response variables are continuous.
2. The residuals follow the multivariate normal probability distribution with mean zero and constant variance-covariance matrix.
3. The subjects are independent.

Technical Details

General Linear Multivariate Model

This section provides the technical details of the MANOVA designs that can be analyzed by *PASS*. The approximate power calculations outlined in Muller, LaVange, Ramey, and Ramey (1992) are used. Using their notation, for N subjects, the usual general linear multivariate model is

$$\underset{(N \times p)}{Y} = \underset{(N \times q \times p)}{XM} + \underset{(N \times p)}{R}$$

Multivariate Analysis of Variance (MANOVA)

where each row of the residual matrix R is distributed as a multivariate normal

$$\text{row}_k(R) \sim N_p(0, \Sigma)$$

Note that p is the number of response variables and q is the number of design variables, Y is the matrix of responses, X is the design matrix, M is the matrix of regression parameters (means), and R is the matrix of residuals.

Hypotheses about various sets of regression parameters are tested using

$$H_0: \Theta = \Theta_0$$

$$CM = \Theta$$

where C is an orthonormal contrast matrix and Θ_0 is a matrix of hypothesized values, usually zeros. Note that C defines contrasts among the factor levels. Tests of the various main effects and interactions may be constructed with suitable choices of C . These tests are based on

$$\hat{M} = (X'X)^{-1} X'Y$$

$$\hat{\Theta} = C\hat{M}$$

$$H = (\hat{\Theta} - \Theta_0)' [C(X'X)^{-1}C']^{-1} (\hat{\Theta} - \Theta_0)$$

$$E = \hat{\Sigma} \cdot (N - r)$$

$$T = H + E$$

where r is the rank of X .

Wilks' Lambda Approximate F Test

The hypothesis $H_0: \Theta = \Theta_0$ may be tested using Wilks' likelihood ratio statistic W . This statistic is computed using

$$W = |ET^{-1}|$$

An F approximation to the distribution of W is given by

$$F_{df_1, df_2} = \frac{\eta / df_1}{(1 - \eta) / df_2}$$

where

$$\lambda = df_1 F_{df_1, df_2}$$

$$\eta = 1 - W^{1/g}$$

$$df_1 = ap$$

$$df_2 = g[(N - r) - (p - a + 1)/2] - (ap - 2)/2$$

$$g = \left(\frac{a^2 p^2 - 4}{a^2 + p^2 - 5} \right)^{\frac{1}{2}}$$

Multivariate Analysis of Variance (MANOVA)

Pillai-Bartlett Trace Approximate F Test

The hypothesis $H_0: \Theta = \Theta_0$ may be tested using the Pillai-Bartlett Trace. This statistic is computed using

$$T_{PB} = tr(HT^{-1})$$

A noncentral F approximation to the distribution of T_{PB} is given by

$$F_{df_1, df_2} = \frac{\eta / df_1}{(1 - \eta) / df_2}$$

where

$$\lambda = df_1 F_{df_1, df_2}$$

$$\eta = \frac{T_{PB}}{s}$$

$$s = \min(a, p)$$

$$df_1 = ap$$

$$df_2 = s[(N - r) - p + s]$$

Hotelling-Lawley Trace Approximate F Test

The hypothesis $H_0: \Theta = \Theta_0$ may be tested using the Hotelling-Lawley Trace. This statistic is computed using

$$T_{HL} = tr(HE^{-1})$$

An F approximation to the distribution of T_{HL} is given by

$$F_{df_1, df_2} = \frac{\eta / df_1}{(1 - \eta) / df_2}$$

where

$$\lambda = df_1 F_{df_1, df_2}$$

$$\eta = \frac{\frac{T_{HL}}{s}}{1 + \frac{T_{HL}}{s}}$$

$$s = \min(a, p)$$

$$df_1 = ap$$

$$df_2 = s[(N - r) - p - 1] + 2$$

Multivariate Analysis of Variance (MANOVA)

M (Mean) Matrix

In the general linear multivariate model presented above, M represents a matrix of regression coefficients. Although other structures and interpretations of M are possible, in this module we assume that the elements of M are the cell means. The rows of M represent the factor categories and the columns of M represent the response variables. (Note that this is just the opposite of the orientation used when entering M into the spreadsheet.)

The q rows of M represent the q groups into which the subjects can be classified. For example, if a design includes three factors with 2, 3, and 4 categories, the matrix M would have $2 \times 3 \times 4 = 24$ rows. That is, $q = 24$.

Consider now an example in which $q = 3$ and $p = 4$. That is, there are three groups into which subjects can be placed. Each subject has four made. The matrix M would appear as follows.

$$M = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} \\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34} \end{bmatrix}$$

For example, the element μ_{12} is the mean of the second response of subjects in the first group. To calculate the power of this design, you would need to specify appropriate values of all twelve means.

C Matrix – Contrasts

The C matrix is comprised of contrasts that are applied to the rows of M . You do not have to specify these contrasts. They are generated for you. You should understand that a different C matrix is generated for each term in the model.

Generating the C Matrix when there are Multiple Between Factors

Generating the C matrix when there is more than one factor is more difficult. We use the method of O'Brien and Kaiser (1985) which we briefly summarize here.

Step 1. Write a complete set of contrasts suitable for testing each factor separately. For example, if you have three factors with 2, 3, and 4 categories, you might use

$$\ddot{C}_{B1} = \begin{bmatrix} -1 & 1 \\ \sqrt{2} & \sqrt{2} \end{bmatrix}, \ddot{C}_{B2} = \begin{bmatrix} -2 & 1 & 1 \\ \sqrt{6} & \sqrt{6} & \sqrt{6} \\ 0 & -1 & 1 \\ & \sqrt{2} & \sqrt{2} \end{bmatrix}, \text{ and } \ddot{C}_{B3} = \begin{bmatrix} -3 & 1 & 1 & 1 \\ \sqrt{12} & \sqrt{12} & \sqrt{12} & \sqrt{12} \\ 0 & -2 & 1 & 1 \\ & \sqrt{6} & \sqrt{6} & \sqrt{6} \\ 0 & 0 & -1 & 1 \\ & & \sqrt{2} & \sqrt{2} \end{bmatrix}.$$

Step 2. Define appropriate J_k matrices corresponding to each factor. These matrices comprised of one row and k columns whose equal element is chosen so that the sum of its elements squared is one. In this example, we use

$$J_2 = \begin{bmatrix} 1 & 1 \\ \sqrt{2} & \sqrt{2} \end{bmatrix}, J_3 = \begin{bmatrix} 1 & 1 & 1 \\ \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix}, J_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \sqrt{4} & \sqrt{4} & \sqrt{4} & \sqrt{4} \end{bmatrix}$$

Multivariate Analysis of Variance (MANOVA)

Step 3. Create the appropriate contrast matrix using a direct (Kronecker) product of either the \ddot{C}_{B_i} matrix if the factor is included in the term or the J_i matrix when the factor is not in the term. Remember that the direct product is formed by multiplying each element of the second matrix by all members of the first matrix. Here is an example

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 3 \end{bmatrix} \\ = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -2 \\ 3 & 4 & 0 & 0 & -3 & -4 \\ 0 & 0 & 2 & 4 & 0 & 0 \\ 0 & 0 & 6 & 8 & 0 & 0 \\ -1 & -2 & 0 & 0 & 3 & 6 \\ -3 & -4 & 0 & 0 & 9 & 12 \end{bmatrix}$$

As an example, we will compute the C matrix suitable for testing factor $B2$

$$C_{B2} = J_2 \otimes \ddot{C}_{B2} \otimes J_4$$

Expanding the direct product results in

$$\begin{aligned} C_{B2} &= J_2 \otimes \ddot{C}_{B2} \otimes J_4 \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{-2}{\sqrt{12}} & \frac{-2}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ 0 & 0 & \frac{-1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} \\ 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} \end{bmatrix} \end{aligned}$$

Similarly, the C matrix suitable for testing interaction $B2B3$ is

$$C_{B2B3} = J_2 \otimes \ddot{C}_{B2} \otimes \ddot{C}_{B3}$$

We leave the expansion of this matrix *PASS*, but we think you have the idea.

Power Calculations

To calculate statistical power, we must determine distribution of the test statistic under the alternative hypothesis which specifies a different value for the regression parameter matrix B . The distribution theory in this case has not been worked out, so approximations must be used. We use the approximations given by Muller and Barton (1989) and Muller, LaVange, Ramey, and Ramey (1992). These approximations state that under the alternative hypothesis, F_U is distributed as a noncentral F random variable with degrees of freedom and noncentrality shown above. The calculation of the power of a particular test may be summarized as follows.

1. Specify values of X , M , Σ , C , and Θ_0 .
2. Determine the critical value using $F_{crit} = FINV(1 - \alpha, df1, df2)$, where $FINV(\)$ is the inverse of the central F distribution and α is the significance level.
3. Compute the noncentrality parameter λ .

Multivariate Analysis of Variance (MANOVA)

4. Compute the power as

$$Power = 1 - NCFPROB(F_{crit}, df1, df2, \lambda)$$

where $NCFPROB()$ is the noncentral F distribution.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tabs. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tabs

The Design tabs contain many of the options that you will be primarily concerned with.

Solve For

Solve For

This option specifies whether you want to solve for power or sample size. This choice controls which options are displayed in the Sample Size section. Note that no plots are generated when you solve for n .

If you select sample size, two more options appear.

Solve For Sample Size 'Based On'

Specify which test statistic you want to use when solving for a sample size.

Solve For Sample Size Using 'Term'

Specify which term to use when searching for a sample size. The power of this term's test will be evaluated as the search is conducted.

All

If you want to use all terms, select 'All.'

Note

Only terms that are active may be selected. If you select a term that is not active, you will be prompted to select another term.

Power and Alpha

Minimum Power

Enter a value for the minimum power to be achieved when solving for the sample size. The resulting sample size is large enough so that the power is greater than this amount. This search is conducted using either all terms or a specific term.

Power is the probability of rejecting the null hypothesis when it is false. Beta, the probability of obtaining a false negative on the statistical test, is equal to 1-power, so specifying power implicitly specifies beta.

Range

Since power is a probability, its valid range is between 0 and 1.

Recommended

Different disciplines have different standards for power. A popular value is 0.9, but 0.8 is also popular.

Multivariate Analysis of Variance (MANOVA)

Alpha for All Terms

Enter a single value of alpha to be used in all statistical tests.

Alpha is the probability of obtaining a false positive on a statistical test. That is, it is the probability of rejecting a true null hypothesis. The null hypothesis is usually that the variance of the means (or effects) is zero.

Range

Since Alpha is a probability, it is bounded by 0 and 1. Commonly, it is between 0.001 and 0.250.

Recommended

Alpha is usually set to 0.05 for two-sided tests such as those considered here.

Design and Effects

Number of Response Variables

Enter the number of response variables in your design. For a true MANOVA, this value must be greater than one. The number of rows in the means matrix must equal this value. If you specify a covariance matrix, the number of columns specified must equal this value.

Number of Factors

Select the number of factors included in the design. Up to three between and within factors may be selected.

Number of Levels

Specify the number of levels (categories) contained in each active factor. A factor must have at least 2 levels. We have arbitrarily set the maximum at 100, but realistically, the maximum is about 10.

Since the total sample size is equal to the product of the number of levels in each factor and the number of observations per group (n), increasing the number of levels of a factor increases the total sample size of the experiment.

Columns Containing the Means

Specify the spreadsheet columns containing a matrix of mean from which the values of σ_m can be computed for each term. To select the spreadsheet columns and enter the means into the spreadsheet, press the icon directly to the right.

In the spreadsheet, the factors are represented across the columns and the response variables are represented down the rows.

The number of columns specified must equal the number of groups. The number of groups is equal to the product of the number of levels of the factors.

The number of rows with data in these columns must equal the number of response variables. For example, suppose you are designing an experiment that is to have two factors (A & B) and four response variables. The four columns of the spreadsheet would be

A1B1 A1B2 A2B1 A2B2

The rows of the spreadsheet would represent

RV1

RV2

RV3

RV4

Multivariate Analysis of Variance (MANOVA)**K's (Multipliers)**

It is often useful to conduct a sensitivity analysis to determine the relationship between the means and the power or sample size. Instead of having to enter all values of the means over and over, this option allows you to enter multipliers that will be applied to the means. A separate power calculation or sample size search is generated for each value of K. These values become the horizontal axis in the second power chart. If you want to ignore this setting, enter '1'.

Sample Size (when 'Solve For' = Power)

The options available for sample size depend on the setting of Solve For. This section is for when you have set Solve For to Power.

Group Allocation

Specify how subjects are allocated to groups.

Equal ($n_1 = n_2 = \dots = n$)

The sample size of every group is n . That is, all group sample sizes are equal. The values of n are entered in the box below.

Unequal (Enter n_1, n_2, \dots Individually)

The sample sizes of each group are possibly different and are entered individually as a list below.

 n (Size Per Group, Equal Group Size)

Enter n , the number of subjects in each group. Sample sizes must be 2 or more.

You can specify a single value or a list. Enter a single value to be used as the sample size of all groups. If you enter a list of values, a separate power analysis is calculated for each value in the list.

 n_1, n_2, \dots (List)

Enter a list of group sample sizes, one per group. The number of groups is equal to the product of the number factor levels. The number of items entered must be equal to the number of groups. If not enough values are entered, the last value is used for remaining groups.

Group Order

The items in the list correspond to the treatment combinations of the factors. The order of the groups is such that the last factor changes the fastest and the first changes the slowest. Here are some examples of how the groups are numbered.

One Factor

If there is one factor 'A' with five levels, group 1 corresponds to A1, group 2 to A2, and so on.

Two Factors

If there are two factors 'A' and 'B' each with two levels, the groups are assigned as follows:

<u>Group</u>	<u>Level</u>
1	A1B1
2	A1B2
3	A2B1
4	A2B2

Multivariate Analysis of Variance (MANOVA)

Three Factors

If there are three factors 'A,' 'B,' and 'C' each with two levels, the groups are assigned as follows:

Group	Level
1	A1B1C1
2	A1B1C2
3	A1B2C1
4	A1B2C2
5	A2B1C1
6	A2B1C2
7	A2B2C1
8	A2B2C2

Sample Size (when 'Solve For' = Sample Size)

The options available for sample size depend on the setting of Solve For. This section is for when you have set Solve For to Sample Size.

Group Allocation

Specify how subjects are allocated to groups during the sample size search.

Equal ($n_1 = n_2 = \dots$)

The search is conducted with all group sample sizes are equal.

Unequal

The search is conducted using the Allocation Pattern given below. This pattern gives the ratio of each sample size to the smallest (or the proportion of the total sample size in each group).

Allocation Pattern

Enter a list of positive numbers, one per group, that represent the relative sample size of each group. This pattern will be used (within rounding) as the sample size search is conducted. Only the relative magnitudes of the values matter when specifying the pattern.

The number of groups is equal to the product of the factor levels. The number of items entered must be equal to the number of groups. If not enough values are entered, the last value is used for remaining groups.

The numbers in the pattern may be interpreted in different ways:

Proportions

Enter the proportions of all subjects in each group. For example, suppose you have one factor with four levels in which twice as many subjects are desired in the second group as in the other groups.

You might enter: 0.2 0.4 0.2 0.2

Which means: 20% in the first, third, and four groups and 40% in the second.

Ratios

Enter the ratio of each group to the smallest value. For example, suppose you have one factor with four levels in which twice as many subjects are desired in the second group.

You might enter: 1 2 1 1

Which means: 20% in the first, third, and four groups and 40% in the second.

Multivariate Analysis of Variance (MANOVA)

Group Order

The items in the list correspond to the treatment combinations of the factors. The order of the groups is such that the last factor changes the fastest and the first changes the slowest. Here are some examples of how the groups are numbered.

One Factor

If there is one factor 'A' with five levels, group 1 corresponds to A1, group 2 to A2, and so on.

Two Factors

If there are two factors 'A' and 'B' each with two levels, the groups are assigned as follows:

Group	Level
1	A1B1
2	A1B2
3	A2B1
4	A2B2

Three Factors

If there are three factors 'A,' 'B,' and 'C' each with two levels, the groups are assigned as follows:

Group	Level
1	A1B1C1
2	A1B1C2
3	A1B2C1
4	A1B2C2
5	A2B1C1
6	A2B1C2
7	A2B2C1
8	A2B2C2

Variance Tab

This tab specifies the variance-covariance matrix.

Variance-Covariance Matrix Specification

Input Type

This option specifies which method will be used to define the variance-covariance matrix.

- **σ and ρ**

This option generates the variance-covariance matrix based on the settings for the standard deviation (σ) and the autocorrelation (ρ).

- **Variance-Covariance Matrix Variables**

When this option is selected, the variance-covariance matrix is read in from the columns of the spreadsheet. This is the most flexible method, but specifying a variance-covariance matrix is tedious. You will usually only use this method when a specific variance-covariance is given to you.

Note that the spreadsheet is shown by selecting the menus: 'Window' and then 'Data'.

Multivariate Analysis of Variance (MANOVA)

- **Non-Constant σ 's and ρ 's in Spreadsheet**

The variance-covariance matrix is read in from the columns of the spreadsheet. The σ 's are on the diagonal and the ρ 's are on the off-diagonal.

Note that the spreadsheet is shown by selecting the menus: "Window" and then "Spreadsheet", or by pressing the spreadsheet icon at the right.

Variance-Covariance Matrix Specification - Input Method = σ and ρ

The variance-covariance matrix is generated using σ^2 on the diagonal and $\rho\sigma^2$ for all off-diagonal elements.

σ (Standard Deviation)

This value is used to generate the variance-covariance matrix. This option specifies a standard deviation to be used for all response variables. The square of this value becomes the diagonal elements of the variance-covariance matrix. Since this is a standard deviation, it must be greater than zero.

ρ (Autocorrelation)

Specify an autocorrelation to be used in calculating the off-diagonal elements of the variance-covariance matrix. Since this is a correlation, it must be between -1 and 1.

Correlation Pattern

This option specifies the pattern of the correlations in the variance-covariance matrix. Two options are available:

- **Constant**

The value of ρ is used as the constant correlation. For example, if $\rho = 0.6$ and $p = 6$, the correlation matrix would appear as

$$R = \begin{bmatrix} 1 & 0.600 & 0.600 & 0.600 & 0.600 & 0.600 \\ 0.600 & 1 & 0.600 & 0.600 & 0.600 & 0.600 \\ 0.600 & 0.600 & 1 & 0.600 & 0.600 & 0.600 \\ 0.600 & 0.600 & 0.600 & 1 & 0.600 & 0.600 \\ 0.600 & 0.600 & 0.600 & 0.600 & 1 & 0.600 \\ 0.600 & 0.600 & 0.600 & 0.600 & 0.600 & 1 \end{bmatrix}$$

- **AR(1)**

The value of ρ is used as the base autocorrelation in a first-order, serial correlation pattern. For example, $\rho = 0.6$ and $p = 6$, the correlation matrix would appear as

$$R = \begin{bmatrix} 1 & 0.600 & 0.360 & 0.216 & 0.130 & 0.078 \\ 0.600 & 1 & 0.600 & 0.360 & 0.216 & 0.130 \\ 0.360 & 0.600 & 1 & 0.600 & 0.360 & 0.216 \\ 0.216 & 0.360 & 0.600 & 1 & 0.600 & 0.360 \\ 0.130 & 0.216 & 0.360 & 0.600 & 1 & 0.600 \\ 0.078 & 0.130 & 0.216 & 0.360 & 0.600 & 1 \end{bmatrix}$$

This pattern is often chosen if the response variables are measured across time.

Multivariate Analysis of Variance (MANOVA)

Covariance Matrix Specification - Input Method = 'Covariance Matrix Variables'

This option instructs the program to read the covariance matrix from the spreadsheet.

Spreadsheet Columns Containing the Variance-Covariance Matrix

This option designates the columns on the current spreadsheet holding the covariance matrix. The number of columns and number of rows must match the number of response variables at which the subjects are measured. The rows of this matrix represent the values corresponding to the response variables.

Reports Tab

This tab specifies which reports and graphs are displayed as well as their format.

Select Output

References

Check this box to display the references.

Definitions

Check this box to display the report definitions.

Means Matrix

Check this box display the Means matrix.

SD and Autocorrelation Matrix

Check this box display the standard deviations and autocorrelations.

Summary Statements

Indicate whether to show the summary statements and which test that is to be reported on in the Summary Statements.

Report Options

Maximum Term-Order Reported

Indicate the maximum order of terms to be reported on. Occasionally, higher-order interactions are of little interest and so they may be omitted. For example, enter a '2' to limit output to individual factors and two-way interactions.

Decimal Places for Numeric Reports

Specify the number of decimal places in the reports for each item.

Multivariate Analysis of Variance (MANOVA)

Example 1 – Determining Power

Researchers are planning a study of the impact of a drug. They want to evaluate the differences in heart rate and blood pressure among three age groups: 20-40, 41-60, and over 60. They want to be able to detect a 10% change in heart rate and in blood pressure among the age groups. They plan to analyze the data using Wilks' lambda.

Previous studies have found an average heart rate of 93 with a standard deviation of 4 and an average blood pressure of 130 with a standard deviation of 5. The correlation between the two responses will be set at 0.7.

From a heart rate of 93, a 10% reduction gives 84. They want to be able to detect age-group heart-rate means the range from 93 to 84. From a blood pressure of 130, a 10% reduction gives 117. They want to be able to detect age-group blood-pressure means that range from 130 to 117. Hence, the means matrix that they will use is

<u>C1</u>	<u>C2</u>	<u>C3</u>
93	88	84
130	124	117

Based on the standard deviation settings that they chose to use, the variance-covariance matrix will be

<u>C4</u>	<u>C5</u>
16	14
14	25

In order to understand the relationship between power and sample size, they decide to calculate power values for sample sizes between 2 and 12, using a 0.05 significance level.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multivariate Analysis of Variance (MANOVA)** procedure window by expanding **Means**, then **Multivariate Means**, then clicking on **MANOVA**, and then clicking on **Multivariate Analysis of Variance (MANOVA)**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**. You can see that the values have been loaded into the spreadsheet by clicking on the spreadsheet button.

<u>Option</u>	<u>Value</u>
Design: Means Tab	
Solve For	Power
Wilks' Lambda	Checked
Alpha for All Terms	0.05
Number of Response Variables.....	2
Number of Factors.....	1
Number of Levels in Factor A	3
Columns Containing the Means	C1-C3
Press the Spreadsheet button to enter the following values into the spreadsheet for columns C1 through C3	
Row 1.....	93 88 84
Row 2.....	130 124 117
K's (Multipliers)	0.5 1 1.5
Group Allocation	Equal (n1 = n2 = ... = n)
n (Size Per Group).....	2 4 6 8 10 12

Multivariate Analysis of Variance (MANOVA)

Design: Variance Tab

Input Type..... Variance-Covariance Matrix in Spreadsheet

Columns Containing Covariance Matrix . C4-C5

Press the Spreadsheet button to enter the following values into the spreadsheet for columns C4 and C5

Row 1..... 16 14

Row 2..... 14 25

Reports Tab

Numeric Results by Term Checked

Numeric Results by Design Not Checked

Means Matrix Checked

Variance-Covariance Matrix Checked

SD and Autocorrelation Matrix..... Checked

Summary Statements Checked

Number of Summary Statements 1

Test Statistic (for Summary) Wilks' Lambda

Plots Tab

Plots of Power vs n by Terms..... Checked

Plots of Power vs K by n..... Checked

Maximum Order of Terms Plotted 3

Test that is Plotted..... Wilks' Lambda

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Term Report

Results for Factor A (Levels = 3)										
Test	Power	n	N	Multiply Means		Test Statistic	Approx. F Statistic	DF1 DF2	Alpha	Beta
				By	Statistic					
Wilks	0.0729	2	6	0.5	0.622	0.27	4 4	0.05	0.93	
Wilks	0.1291	2	6	1.0	0.286	0.87	4 4	0.05	0.87	
Wilks	0.2046	2	6	1.5	0.146	1.62	4 4	0.05	0.80	
Wilks	0.1888	4	12	0.5	0.712	0.74	4 16	0.05	0.81	
Wilks	0.5749	4	12	1.0	0.378	2.51	4 16	0.05	0.43	
Wilks	0.8722	4	12	1.5	0.208	4.78	4 16	0.05	0.13	
Wilks	0.3191	6	18	0.5	0.733	1.17	4 28	0.05	0.68	
Wilks	0.8548	6	18	1.0	0.403	4.02	4 28	0.05	0.15	
Wilks	0.9916	6	18	1.5	0.226	7.71	4 28	0.05	0.01	
Wilks	0.4488	8	24	0.5	0.743	1.60	4 40	0.05	0.55	
Wilks	0.9603	8	24	1.0	0.415	5.51	4 40	0.05	0.04	
Wilks	0.9997	8	24	1.5	0.236	10.61	4 40	0.05	0.00	
Wilks	0.5678	10	30	0.5	0.748	2.03	4 52	0.05	0.43	
Wilks	0.9907	10	30	1.0	0.422	7.00	4 52	0.05	0.01	
Wilks	1.0000	10	30	1.5	0.241	13.49	4 52	0.05	0.00	
Wilks	0.6704	12	36	0.5	0.752	2.46	4 64	0.05	0.33	
Wilks	0.9981	12	36	1.0	0.427	8.48	4 64	0.05	0.00	
Wilks	1.0000	12	36	1.5	0.244	16.37	4 64	0.05	0.00	

Summary Statements
A MANOVA design with 1 factor and 2 response variables has 3 groups with 2 subjects each for a total of 6 subjects. This design achieves 7% power to test factor A if a Wilks' Lambda Approximate F Test is used with a 5% significance level.

This report gives the power for each value of n and K . It is useful when you want to compare the powers of the terms in the design at a specific sample size.

Multivariate Analysis of Variance (MANOVA)

In this example, for $K = 1$, the design goal of 0.95 power is achieved for $n = 8$.

The definitions of each of the columns of the report are as follows.

Test

This column identifies the test statistic. Since the power depends on the test statistic, you should make sure that this is the test statistic that you will use in your analysis.

Power

This is the computed power for the term.

n

The value of n is the number of subjects per group.

N

The value of N is the total number of subjects in the study.

Multiply Means By

This is the value of the means multiplier, K .

Test Statistic

This is the value of the test statistic computed at the hypothesized values. The name of the statistic is identified in the Test column. Possible values are Wilks' lambda, Pillai-Bartlett trace, or Hotelling-Lawley trace. The actual formulas used were given earlier in the Technical Details section.

Approx. F Statistic

This is the value of the F statistic that is used to compute the probability levels. This value is calculated using the hypothesized values. The actual formulas used were given earlier in the Technical Details section.

DF1|DF2

These are the numerator and denominator degrees of freedom of the approximating F distribution.

Alpha

Alpha is the significance level of the test.

Beta

Beta is the probability of failing to reject the null hypothesis when the alternative hypothesis is true.

Means Matrix

Means Matrix Section			
Name	A1	A2	A3
Y1	93.00	88.00	84.00
Y2	130.00	124.00	117.00

This report shows the means matrix that was read in. It may be used to get an impression of the magnitude of the difference among the means that is being studied. When a Means Multiplier, K , is used, each value of K is multiplied times each value of this matrix.

Multivariate Analysis of Variance (MANOVA)

Variance-Covariance Matrix Section

Variance-Covariance Matrix Section

Response	Y1	Y2
Y1	16.000	14.000
Y2	14.000	25.000

This report shows the variance-covariance matrix that was read in from the spreadsheet or generated by the settings of on the Covariance tab. The standard deviations are given on the diagonal and the correlations are given off the diagonal.

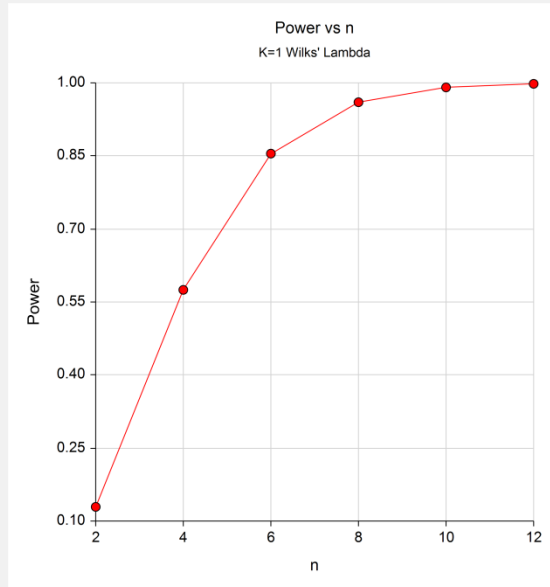
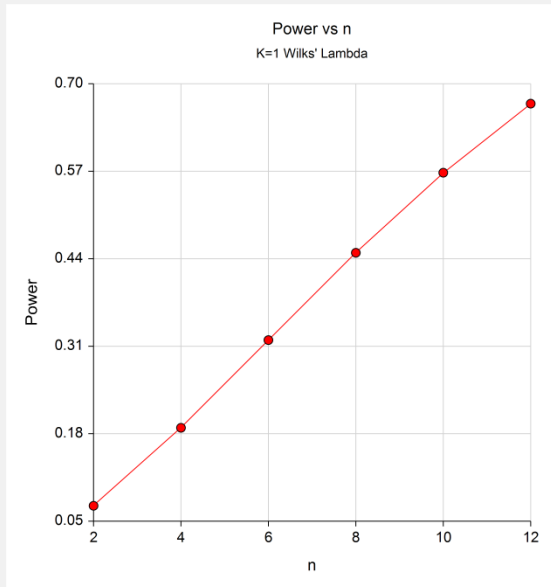
Standard Deviations and Autocorrelations Section

Standard Deviations and Autocorrelations Section

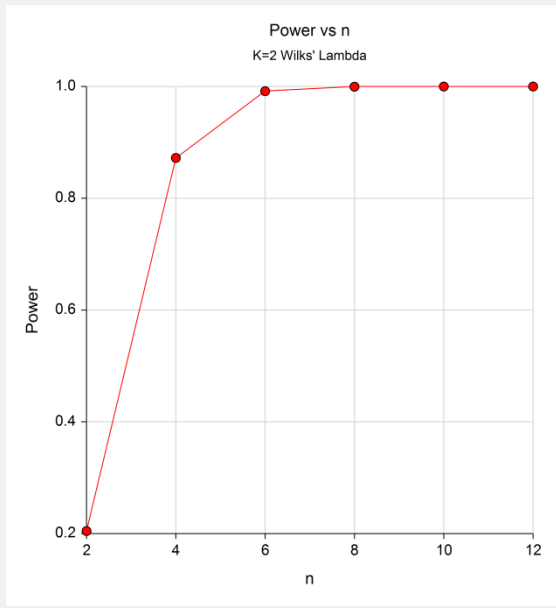
Response	Y1	Y2
Y1	4.000	0.700
Y2	0.700	5.000

This report shows the variance-covariance matrix that was read in from the spreadsheet or generated by the settings of on the Covariance tab. The standard deviations are given on the diagonal and the correlations are given off the diagonal.

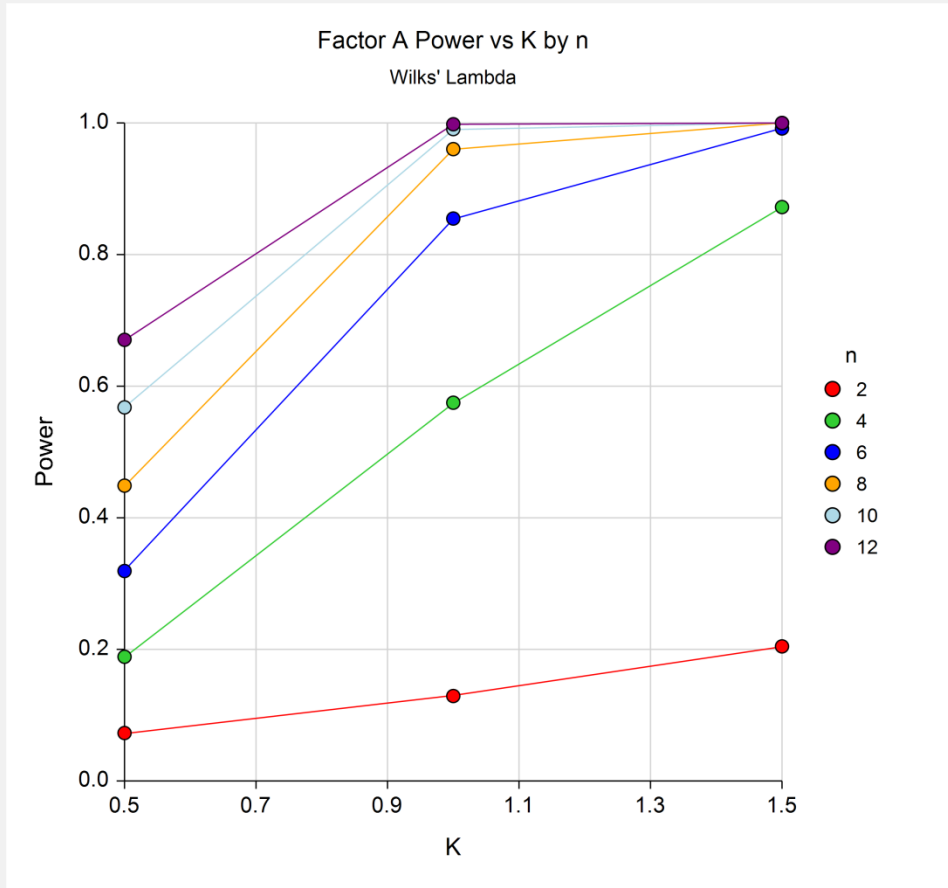
Plots Section



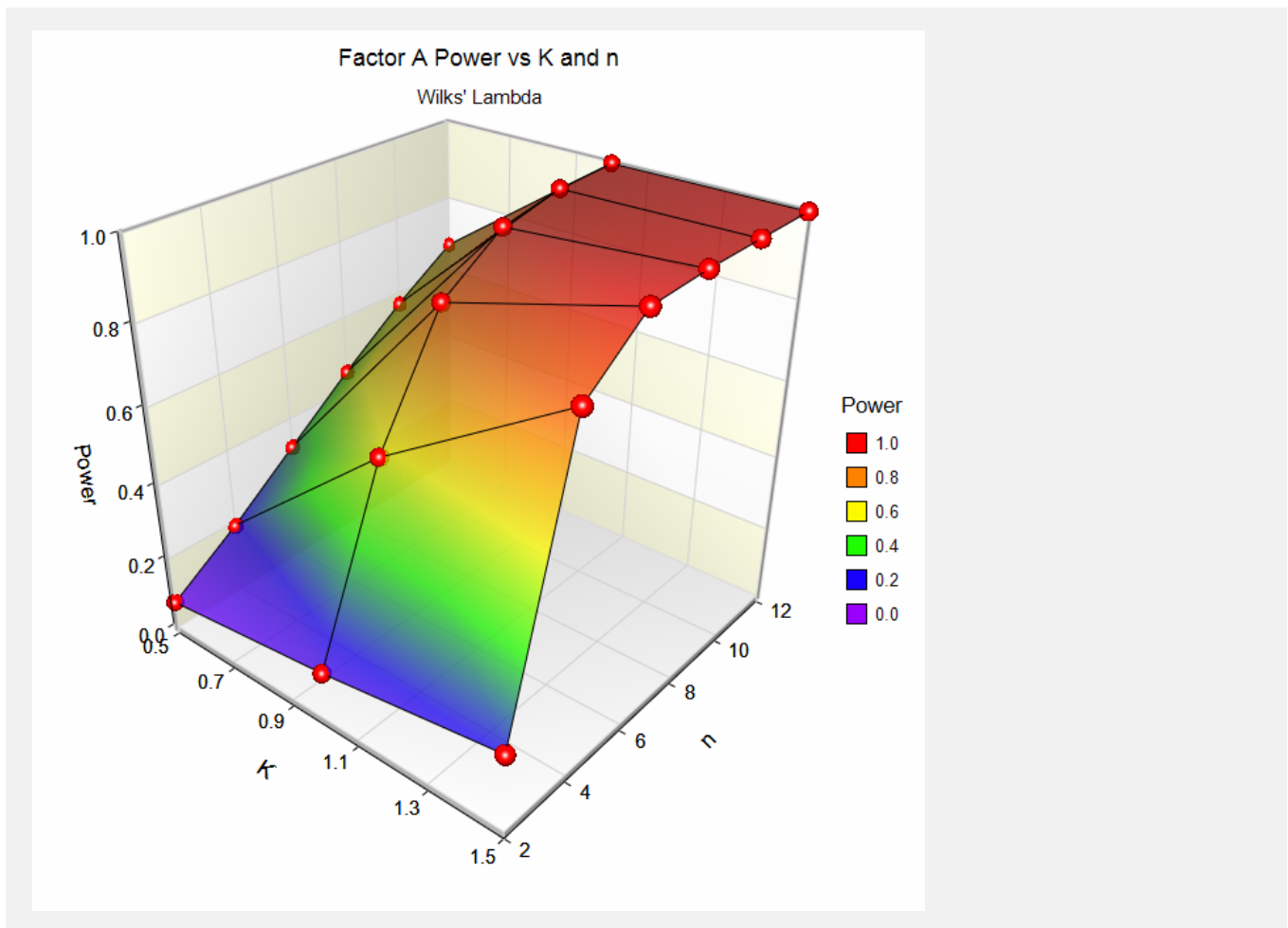
Multivariate Analysis of Variance (MANOVA)



These charts show the relationship between power and n for each value of K . Note that high-order interactions may be omitted from the plot by reducing the Maximum Term-Order Plotted option on the Reports tab.



Multivariate Analysis of Variance (MANOVA)



These charts show the relationship between power and K for each value of n . Remember that K is the mean multiplier. It changes the effect size.

Example 2 – Validation

In this example, we will set $p = 2$, $q = 3$, $\alpha = 0.05$, and $n = 4$. The mean and covariance matrices are

$$M = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$$

The contrast matrix C is

$$C = \begin{bmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

The $X'X$ matrix is

$$X'X = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

The matrix Θ is

$$\begin{aligned} \Theta &= CM \\ &= \begin{bmatrix} \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

The matrix H is

$$\begin{aligned} H &= (\hat{\Theta} - \Theta_0)' [C(X'X)^{-1}C']^{-1} (\hat{\Theta} - \Theta_0) \\ &= \begin{bmatrix} 8 & 4 \\ 4 & 8/3 \end{bmatrix} \end{aligned}$$

The matrix E is

$$\begin{aligned} E &= \hat{\Sigma} \cdot (N - r) \\ &= \begin{bmatrix} 36 & 9 \\ 9 & 36 \end{bmatrix} \end{aligned}$$

The matrix T is

$$\begin{aligned} T &= H + E \\ &= \begin{bmatrix} 44 & 13 \\ 13 & 38\frac{2}{3} \end{bmatrix} \end{aligned}$$

Multivariate Analysis of Variance (MANOVA)

Using these matrices, we can calculate the values of the test statistics. We will only calculate the results for Wilks' lambda. We have

$$\begin{aligned} W &= \det(ET^{-1}) \\ &= 0.79290842 \end{aligned}$$

$$\begin{aligned} a &= q - 1 \\ &= 2 \end{aligned}$$

$$\begin{aligned} g &= \left(\frac{a^2 p^2 - 4}{a^2 + p^2 - 5} \right)^{\frac{1}{2}} \\ &= \left(\frac{2^2 2^2 - 4}{2^2 + 2^2 - 5} \right)^{\frac{1}{2}} \\ &= 2 \end{aligned}$$

$$\begin{aligned} \eta &= 1 - W^{1/g} \\ &= 1 - \sqrt{0.79290842} \\ &= 0.10954595 \end{aligned}$$

$$\begin{aligned} df_1 &= ap \\ &= 4 \end{aligned}$$

$$\begin{aligned} df_2 &= g[(N - r) - (p - a + 1) / 2] - (ap - 2) / 2 \\ &= 2[(12 - 3) - (2 - 2 + 1) / 2] - (4 - 2) / 2 \\ &= 16 \end{aligned}$$

$$\begin{aligned} F_{df_1, df_2} &= \frac{\eta / df_1}{(1 - \eta) / df_2} \\ &= \frac{0.10954595 / 4}{(1 - 0.10954595) / 16} \\ &= 0.49209030 \end{aligned}$$

$$\begin{aligned} \lambda &= df_1 F_{df_1, df_2} \\ &= 4(0.49209030) \\ &= 1.96836120 \end{aligned}$$

For an F with 4 and 16 degrees of freedom, the 5% critical value is 3.0069172799. Finally, compute the power using a noncentral F with 4 and 16 degrees of freedom and noncentrality parameter

$$\begin{aligned} Power &= \Pr(f > F | df_1 = 4, df_2 = 16, \lambda = 1.96836120) \\ &= 0.1370631884 \end{aligned}$$

Multivariate Analysis of Variance (MANOVA)

In order to run this example in *PASS*, the values of the means and the covariance matrix (given above) must be entered on a spreadsheet. The instructions below assume that the means are in columns C1-C3, while the covariance matrix is in columns C4-C5.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multivariate Analysis of Variance (MANOVA)** procedure window by expanding **Means**, then **Multivariate Means**, then clicking on **MANOVA**, and then clicking on **Multivariate Analysis of Variance (MANOVA)**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**. You can see that the values have been loaded into the spreadsheet by clicking on the spreadsheet button.

Option Value

Design: Means Tab

Solve For **Power**
 Wilks' Lambda **Checked**
 Alpha for All Terms **0.05**
 Number of Response Variables..... **2**
 Number of Factors..... **1**
 Number of Levels in Factor A **3**
 Columns Containing the Means **C1-C3**
 Press the Spreadsheet button to enter the following values into the spreadsheet for columns C1 through C3

Row 1.....	1	2	3
Row 2.....	1	1	2

K's (Multipliers) **1**
 Group Allocation **Equal (n1 = n2 = ... = n)**
 n (Size Per Group)..... **4**

Design: Variance Tab

Input Type..... **Variance-Covariance Matrix in Spreadsheet**
 Columns Containing Covariance Matrix . **C4-C5**
 Press the Spreadsheet button to enter the following values into the spreadsheet for columns C4 and C5

Row 1.....	4	1
Row 2.....	1	4

Reports Tab

Numeric Results by Term **Checked**

Output

Click the Calculate button to perform the calculations and generate the following output.

Term Report

Results for Factor A (Levels = 3)									
Test	Power	n	N	Multiply Means By	Test Statistic	Approx. F Statistic	DF1 DF2	Alpha	Beta
Wilks	0.1371	4	12	1.0000	0.793	0.49	4 16	0.0500	0.86

As you can see, the power computed here matches the results we computed manually.