Chapter 544

One-Way Analysis of Variance Contrasts

Introduction

The one-way (multiple group) design allows the means of two or more populations (groups) to be compared to determine if at least one mean is different from the others. The *F* test is used to determine statistical significance.

The usual F-test tests the hypothesis that all means are equal versus the alternative that at least one mean is different from the rest. Often, a more specific alternative is desired. For example, you might want to test whether the treatment means are different from the control mean, the low dose is different from the high dose, a linear trend exists across dose levels, and so on. These questions are tested using specific contrasts.

A *comparison* is a weighted average of the means, in which the weights may be negative. When the weights sum to zero, the comparison is called a *contrast*. **PASS** provides results for contrasts. To specify a contrast, we need only specify the weights.

For example, suppose an experiment conducted to study a drug will have three dose levels: none (control), 20 mg, and 40 mg. The first question is whether the drug made a difference. If it did, the average response for the two groups receiving the drug should be different from the control. If we label the group means M0, M2, and M4, we are interested in comparing M0 with M2 and M4. This can be done in two ways. One way is to construct two tests, one comparing M0 and M2 and the other comparing M0 and M4. Another method is to perform one test comparing M0 with the average of M2 and M4. These tests are conducted using contrasts. The coefficients are as follows:

M0 vs. M2

To compare M0 versus M2, use the coefficients -1, 1, 0. When applied to the group means, these coefficients result in the comparison M0(-1) + M2(1) + M4(0) which reduces to M2-M0. That is, this contrast results in the difference between two group means. We can test whether this difference is non-zero using the *t* test (or *F* test since the square of the *t* test is an *F* test).

M0 vs. M4

To compare M0 versus M4, use the coefficients -1, 0, 1. When applied to the group means, these coefficients result in the comparison M0(-1) + M2(0) + M4(1) which reduces to M4 - M0. That is, this contrast results in the difference between the two group means.

M0 vs. Average of M2 and M4

To compare M0 versus the average of M2 and M4, use the coefficients -2, 1, 1. When applied to the group means, these coefficients result in the comparison M0(-2) + M2(1) + M4(1) which is equivalent to M4 + M2 - 2(M0).

Assumptions

Using the *F* test requires 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 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 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 within the groups 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.

Technical Details for One-Way ANOVA Contrasts

Suppose *G* groups each have a normal distribution and equal means $(\mu_1 = \mu_2 = \cdots = \mu_G)$. Let $N_1 = N_2 = \cdots = N_G$ denote the number of subjects in each group and let *N* denote the total sample size of all groups. Let μ denote the weighted mean of all groups. That is

$$\mu = \sum_{i=1}^{G} \left(\frac{N_i}{N}\right) \mu_i$$

Let σ denote the common standard deviation of all groups.

Suppose you want to test whether the contrast *C*

$$C = \sum_{i=1}^{G} c_i \mu_i$$

is significantly different from zero. Here the c_i's are the contrast coefficients.

Define

$$\sigma_C = \left|\sum_{i=1}^{G} c_i \mu_i\right| \Big/ \sqrt{N \sum_{i=1}^{G} \frac{c_i^2}{N_i}}$$

Define the noncentrality parameter λ_C, as

$$\lambda_C = N\sigma_C^2 / \sigma^2$$

Power Calculations for Contrasts

The calculation of the power of a test proceeds as follows:

1. Determine the critical value, $F_{1,N-G,\alpha}$, where α is the probability of a type-I error and *G* and *N* are defined above. Note that this is a two-tailed test as no direction is assigned in the alternative hypothesis.
2. From a hypothesized set of μ_i's, calculate the noncentrality parameter λ_C.
3. Compute the power as the probability of being greater than $F_{1,N-G,\alpha}$ on a noncentral-*F* distribution with noncentrality parameter λ_C.

Procedure Options

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

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. Select either *Power* for a power analysis or *Sample Size* for a sample size determination.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of a zero contrast value when in fact the contrast value is not zero.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of a zero contrast value when the contrast value is zero.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size and Group Allocation

G (Number of Groups)

This is the number of groups (arms) whose means are being compared. The number of items used in the Group Allocation boxes is controlled by this number.

This value must be an integer greater than or equal to two.

Group Allocation Input Type (when Solve For = Power or Effect Size)

Specify how you want to enter the information about how the subjects are allocated to each of the G groups.

Possible options are:

- **Equal (N1 = N2 = ... = NG)**
  The sample size of all groups is Ni. Enter one or more values for the common group sample size.

- **Enter group multipliers**
  Enter a list of group multipliers (r1, r2, ..., rG) and one or more values of Ni. The individual group sample sizes are found by multiplying the multipliers by Ni. For example, N1 = r1 x Ni.

- **Enter N1, N2, ..., NG**
  Enter a list of group sample sizes, one for each group.

- **Enter columns of Ni's**
  Select one or more columns of the spreadsheet that each contain a set of group sample sizes going down the column. Each column is analyzed separately.

Ni (Subjects Per Group)

Enter *Ni*, the number of subjects in each group. The total sample size, *N*, is equal to *Ni* x G.

You can specify a single value or a list.

Single Value

Enter a value for the individual sample size of all groups. If you enter '10' here and there are five groups, then each group will be assigned 10 subjects and the total sample size will be 50.

List of Values

A separate power analysis is calculated for each value of Ni in the list. All analyses assume that the common, group sample size is Ni.

Range of *Ni*

Ni > 1

Group Multipliers (r1, r2, ..., rG)

Enter a set of G multipliers, one for each group.

The individual group sample sizes is computed as *Ng* = ceiling[*rg* x *Ni*], where ceiling[y] is the first integer greater than or equal to *y*. For example, the multipliers {1, 1, 2, 2.95} and *base Ni* of 10 would result in the sample sizes {10, 10, 20, 30}.

Incomplete List

If the number of items in the list is less than *G*, the missing multipliers are set equal to the last entry in the list.

Range

The items in the list must be positive. The resulting sample sizes must be at least 1.

Ni (Base Subjects Per Group)

Enter *Ni*, the base sample size of each group. The number of subjects in the group is found by multiplying this number by the corresponding group multiplier, {*r1*, *r2*, ..., *rG*}, and rounding up to the next integer.

You can specify a single value or a list.

Single Value

Enter a value for the base group subject count.

List of Values

A separate power analysis is calculated for each value of *Ni* in the list.

Range

Ceiling[*Ni* x *ri*] ≥ 1.

N1, N2, …, NG (List)

Enter a list of G subject counts, one for each group.

Incomplete List

If the number of items in the list is less than G, the missing subject counts are set equal to the last entry in the list.

Range

The items in the list must be positive. At least one item in the list must be greater than 1.

Columns of Ni's

Enter one or more spreadsheet columns containing vertical lists of group subject counts.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

On the spreadsheet, the group subject counts are entered going down.

Examples (assuming G = 3)

C1	C2	C3
111	1	28
115	20	68
100	30	46

Definition of a Single Column

Each column gives one list. Each column results in a new scenario. The columns are not connected, but all should have exactly G rows.

Each entry in the list is the subject count of that group.

Incomplete List

If the number of items in the list is less than G, the missing entries are set equal to the last entry in the list.

Valid Entries

All values should be positive integers. At least one value must be greater than one.

Note

The column names (C1, C2, ...) can be changed by right-clicking on them in the spreadsheet.

Group Allocation Input Type (when Solve For = Sample Size)

Specify how you want to enter the information about how the subjects are allocated to each of the G groups.

Options

- **Equal (N1 = N2 = ... = NG)**

  All group subject counts are equal to Ni. The value of Ni will be found by conducting a search.

- **Enter group allocation pattern**

  Enter an allocation pattern (r1, r2, ..., rG). The pattern consists of a set of G numbers. These numbers will be rescaled into proportions by dividing each item by the sum of all items. The individual group subject counts are found by multiplying these proportions by N (the total subject count) and rounding up.

- **Enter columns of allocation patterns**

  Select one or more columns of the spreadsheet that each contain a group allocation pattern going down the column. Each column is analyzed separately.

Group Allocation Pattern (r1, r2, ..., rG)

Enter an allocation pattern (*r1*, *r2*, ..., *rG*). The pattern consists of a set of *G* numbers. These numbers will be rescaled into proportions of N by dividing each item by the sum of all items. The individual group subject counts are found by multiplying these proportions by *N* (the total subject count) and rounding up.

For example, the pattern {1, 3, 4} will be rescaled to {0.125, 0.375, 0.5}. The group subject counts will be constrained to these proportions (within rounding) during the search for the subject count configuration that meets the power requirement.

Incomplete List

If the number of items in the list is less than *G*, the missing numbers are set equal to the last entry in the list.

Range

The items in the list must be positive. The resulting subject counts must be at least 1.

Columns of Group Allocation Patterns

Enter one or more spreadsheet columns containing vertical lists of group allocation patterns.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

On the spreadsheet, the group allocation patterns are entered going down.

Examples (assuming G = 3)

C1	C2	C3
1	1	3
1	2	3
3	2	1

Definition of a Single Column

Each column gives one allocation pattern. Each column results in a new scenario. The columns are not connected, but all should have exactly G rows.

Incomplete List

If the number of items in a list is less than G, the missing numbers are set equal to the last entry in the list before they are rescaled.

Valid Entries

All values should be positive numbers. You can enter decimal values.

Note

The column names (C1, C2, ...) can be changed by right-clicking on them in the spreadsheet.

Effect Size

μi's Input Type

Specify how you want to enter the G group means μ1, μ2, ..., μG assumed by the alternative hypothesis. The power is calculated for these values.

Note that under the null hypothesis, these means are all equal.

Options

- **μ1, μ2, ..., μG**

  Specify the values of the group means. These are combined with the contrast coefficients to form σc which is used in the power analysis. The formula for this is

  $$\sigma c = |\Sigma(ci\mu i)| / \sqrt{[N\ \Sigma(ci^2/ni)]}$$

- **μ1, μ2, ..., μG and Multipliers**

  Specify the values of the group means as well as one or more multipliers for quickly generating sets of means.

- **Columns Containing Sets of μi's**

  Select one or more columns of the spreadsheet that each contain a set of μi's going down the column. Each column is analyzed separately.

μ1, μ2, ..., μG

Enter the values of the G group means under the alternative hypothesis. The mean for a particular group is the average response of all subjects in that group.

Range

Each μi should be numeric and at least one of the values must be different from the rest. Also, the result of applying a contrast cannot be zero.

Example

10 10 10 40

Incomplete List

If the number of items in a list is less than G, the missing numbers are set equal to the last entry.

K (Means Multiplier)

Enter one or more values for K, the means multiplier. A separate power calculation is conducted for each value of K. In each analysis, all means (μi's) are multiplied by K. In this way, you can determine how sensitive the power values are to the magnitude of the means without the need to change them individually.

For example, if the original means are '0 1 2', setting this option to '1 2' results in two sets of means used in separate analyses: '0 1 2' in the first analysis and '0 2 4' in the second analysis.

Examples

1

0.5 1 1.5

0.8 to 1.2 by 0.1

Columns Containing Sets of μi's

Enter one or more spreadsheet columns containing vertical lists of μ1, μ2, ..., μG.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

On the spreadsheet, the μi's are entered going down.

Examples (assuming G = 3)

C1	C2	C3
10	10	30
10	20	30
30	20	10

Definition of a Single Column

Each column gives one set of means. Each column results in a new scenario. The columns are not connected, but all should have exactly G rows.

Incomplete List

If the number of items in a list is less than G, the missing numbers are set equal to the last entry in the list.

Valid Entries

You can enter any numeric value.

Note

The column names (C1, C2, ...) can be changed by clicking on them in the spreadsheet.

Contrast Input Type

Specify how you want to enter the contrast coefficients.

A contrast is a weighted average of the group means in which the weights sum to zero. For example, suppose you are studying four groups and that the main hypothesis of interest is whether there is a linear trend across the groups. You might enter "-3, -1, 1, 3" here. This would form the weighted average of the means:

-3 μ1 - μ2 + μ3 + 3 μ4

The point to realize is that these weights (the coefficients) are used to calculate a specific weighted average of the means which is to be compared against zero using a standard, one-degree-of-freedom F-test. Hence, G coefficients (weights) must be defined for each contrast.

These coefficients must sum to zero. Also, the scale of the coefficients does not matter. That is the power of "0.5 0.25 0.25", "-2 1 1", and "-200 100 100" are all the same.

Options

- **List of Contrast Coefficients**

  Enter a list of coefficients, separated by commas or blanks.

- **Multiple Lists of Contrast Coefficients**

  Enter several contrasts, one per column, on the spreadsheet. Each contrast is entered down a column, one coefficient per row.

- **First Group vs Rest**

  A contrast is generated appropriate for testing that the mean of the first group is different from the average of the remaining groups. For example, if there were four groups, the generated coefficients would be "-3, 1, 1, 1".

- **Last Group vs Rest**

  A contrast is generated appropriate for testing that the log mean of the last group is different from the average log value of the remaining groups. For example, if there were four groups, the generated coefficients would be "1, 1, 1, -3".

- **Linear Trend**

  A set of coefficients is generated appropriate for testing the alternative hypothesis that there is a linear (straight-line) trend across the group means. These coefficients assume that the groups are equally spaced across some unspecified, quantitative variable associated with the groups.

- **Maximum Power**

  A set of coefficients is generated which will result in the maximum possible power. This contrast is based on a knowledge of the actual population means, so in practice it cannot be attained. However, it lets you determine how close your power is compared with the maximum possible.

  The reference for the maximum power is given in Winer (1991), page 151. The formula for the contrast coefficients is

$$c_j = N_j\left(\mu_j - \mu\right) / \sqrt{\sum_{j=1}^{G} N_j\left(\mu_j - \mu\right)^2}$$

(List of) Contrast Coefficients

Specify how you want to enter the contrast coefficients.

Enter a list of numbers, separated by blanks or commas, one number for each group. These numbers must sum to zero. The attained power is highly related to the specific contrast that is entered.

Incomplete List

If the list is not long enough, the list is padded with zeros.

Examples

-1 -1 2

-2 -1 0 1 2

-3 -1 1 3

-3 1 1 1

Multiple Lists of (Contrast) Coefficients

Select a set of columns on the spreadsheet which each contain a contrast defined vertically.

Press the Spreadsheet icon (directly to the right) to select the columns and then enter the values.

Press the Input Spreadsheet icon (to the right and slightly up) to view/edit the spreadsheet. Also note that you can obtain the spreadsheet by selecting "Tools", then "Input Spreadsheet", from the menus.

The attained power is highly related to the specific contrast that is entered.

Incomplete List

If the list is not long enough, it is padded with zeros.

Examples

C1 - C3

C1, C2, C3

Con1, Con2, Con3

Spreadsheet Example

If G = 4 and there are three columns specified, the spreadsheet might look as follows.

C1	C2	C3
1	-3	1
1	-1	-1
1	1	-1
-3	3	1

σ (Standard Deviation)

This is σ, the standard deviation between subjects within a group. It represents the variability from subject to subject that occurs when the subjects are treated identically. It is assumed to be the same for all groups. This value is approximated in an analysis of variance table by the square root of the mean square error.

Since they are positive square roots, the numbers must be strictly greater than zero. You can press the σ button to obtain further help on estimating the standard deviation.

You can enter a single value such as ‘10’ or a series of values such as ‘10 20 30 40 50’ or ‘1 to 5 by 0.5’.

Example 1 – Finding Power

An experiment is being designed to compare the means of four groups using a contrast test with a significance level of 0.05. The first group is a control group. The other three groups will have slightly different treatments applied. The researchers are mainly interested in whether the three treatment groups are different from the control group. Hence, they want to test the contrast represented by the coefficients -3, 1, 1, 1. Treatment means of 40, 10, 10, and 10 represent clinically important group differences.

Previous studies have had standard deviations between 18 and 24. To better understand the relationship between power and sample size, the researcher wants to compute the power for several group sample sizes between 2 and 14. The sample sizes will be equal across all groups.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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**.

Option	Value
Design Tab	
Solve For	**Power**
Alpha	**0.05**
G (Number of Groups)	**4**
Group Allocation Input Type	**Equal (N1 = N2 = ⋯ = NG)**
Ni (Subjects Per Group)	**2 4 6 8 10 12 14**
μi's Input Type	**μ1, μ2, ..., μG**
μ1, μ2, ..., μG	**40 10 10 10**
Contrast Input Type	**List of Contrast Coefficients**
Contrast Coefficients	**-3 1 1 1**
σ (Standard Deviation)	**18 21 24**

Annotated Output

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

Numeric Results

Numeric Results
Number of Groups: 4

Power	Total Sample Size N	Subjects Per Group Ni	Group Means Set μi	Cont Coef Set Ci	Std Dev σ	Effect Size σc/σ	Alpha
0.3471	8	2	μi(1)	C(1)	18.00	0.722	0.050
0.2713	8	2	μi(1)	C(1)	21.00	0.619	0.050
0.2201	8	2	μi(1)	C(1)	24.00	0.541	0.050
0.7550	16	4	μi(1)	C(1)	18.00	0.722	0.050
0.6231	16	4	μi(1)	C(1)	21.00	0.619	0.050
0.5123	16	4	μi(1)	C(1)	24.00	0.541	0.050
0.9194	24	6	μi(1)	C(1)	18.00	0.722	0.050
0.8218	24	6	μi(1)	C(1)	21.00	0.619	0.050
0.7132	24	6	μi(1)	C(1)	24.00	0.541	0.050
0.9761	32	8	μi(1)	C(1)	18.00	0.722	0.050
0.9218	32	8	μi(1)	C(1)	21.00	0.619	0.050

(Report Continues)

Set(Set Number): Values
μi(1): 40.00, 10.00, 10.00, 10.00
C(1): -3.000, 1.000, 1.000, 1.000

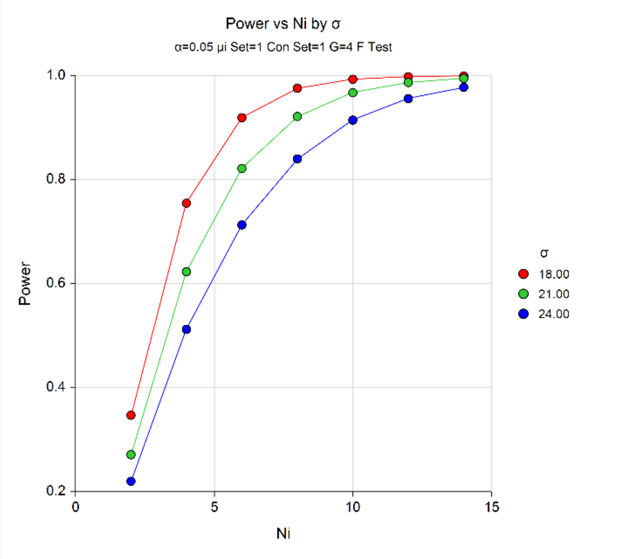
References
Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.
Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.
Kirk, Roger E. 1982. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole. Pacific Grove, California.

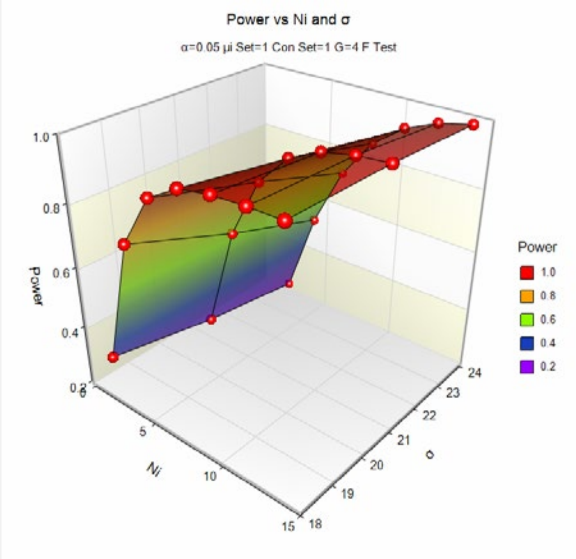
Report Definitions
Power is the probability of rejecting a false null hypothesis.
Total Sample Size N is the total number of subjects in the study.
Subjects Per Group Ni is the number of subjects per group.
Group Means Set μi gives the name and number of the set containing the mean responses for each group.
Cont Coef Set Ci gives the name of the set containing the contrast coefficients that are combined with the group response probabilities.
Std Dev σ is the common standard deviation of the responses within a group.
Effect Size σc/σ is a measure of the effect size. It is the ratio of σc (the SD of the contrast of the means) and σ.
Alpha is the significance level of the test: the probability of rejecting the null hypothesis of equal means when it is true.

Summary Statements
In a one-way ANOVA study, a sample of 8 subjects, divided among 4 groups, achieves a power of 0.3471. This power assumes an F test is used to test a planned comparison (contrast) with contrast coefficients -3.000, 1.000, 1.000, 1.000 at a significance level of 0.050. The group subject counts are 2, 2, 2, 2. The group means under the alternative hypothesis are 40.00, 10.00, 10.00, 10.00. The common standard deviation of the responses is 18.00. The effect size is 0.722.

This report shows the numeric results of this power study.

Plots Section





These plots give a visual presentation to the results in the Numeric Report. We can quickly see the impact on the power of increasing the sample size and the increase in the significance level.

When you create one of these plots, it is important to use trial and error to find an appropriate range for the horizontal variable so that you have results with both low and high power.

Example 2 – Validation using Hand Calculations

We will compute the following example by hand and then compare that with the results that **PASS** obtains. Here are the settings:

Alpha	0.05
G	3
Allocation	Equal
n	5
Means	1, 2, 3
K	1
Coefficients	-2, 1, 1
σ	5

Using these values, we find the following

C'μ	3
σ_c^2	9/18 = 0.5
λ_C	15 x 0.5/(25) = 0.3
$F_{0.95,1,12}$	4.747225

Power = 0.0797

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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**.

Option	Value
Design Tab	
Solve For	**Power**
Alpha	**0.05**
G (Number of Groups)	**3**
Group Allocation Input Type	**Equal (N1 = N2 = ⋯ = NG)**
Ni (Subjects Per Group)	**5**
μi's Input Type	**μ1, μ2, ..., μG**
μ1, μ2, ..., μG	**1 2 3**
Contrast Input Type	**List of Contrast Coefficients**
Contrast Coefficients	**-2 1 1**
σ (Standard Deviation)	**5**

Output

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

Numeric Results

Numeric Results
Number of Groups: 3

Power	Total Sample Size N	Subjects Per Group Ni	Group Means Set μi	Cont Coef Set Ci	Std Dev σ	Effect Size σc/σ	Alpha
0.0797	15	5	μi(1)	C(1)	5.00	0.141	0.050

Set(Set Number): Values
μi(1): 1.00, 2.00, 3.00
C(1): -2.000, 1.000, 1.000

PASS has also calculated the power to be 0.0797.