Chapter 570

Repeated Measures Analysis

Introduction

This module calculates the power for *repeated measures* designs having up to three between factors and up to three within factors. It computes power for both the univariate (F-test and F-test with Geisser-Greenhouse correction) and multivariate (Wilks' lambda, Pillai-Bartlett trace, and Hotelling-Lawley trace) approaches. It can also be used to calculate the power of *crossover* designs.

Repeated measures designs are popular because they allow a subject to serve as their own control. This usually improves the precision of the experiment. However, when the analysis of the data uses the traditional *F*-tests, additional assumptions concerning the structure of the error variance must be made. When these assumptions do not hold, the Geisser-Greenhouse correction provides reasonable adjustments so that significance levels are accurate.

An alternative to using the *F*-test with repeated measures designs is to use one of the multivariate tests: Wilks' lambda, Pillai-Bartlett trace, or Hotelling-Lawley trace. These alternatives are appealing because they do not make the strict, often unrealistic, assumptions about the structure of the variance-covariance matrix. Unfortunately, they may have less power than the *F*-test and they cannot be used in all situations.

An example of a two-factor repeated measures design that can be analyzed by this procedure is shown by the following diagram.

Group 1			Group 2	
Subject 1	**Subject 2**	**Month**	**Subject 3**	**Subject 4**
Treatment L	Treatment L	1	Treatment L	Treatment L
Treatment M	Treatment M	2	Treatment M	Treatment M
Treatment H	Treatment H	3	Treatment H	Treatment H

Groups 1 and 2 form the *between* factor. The within factor has three levels: *L*, *M*, and *H* (low, medium, and high). There are four subjects in this experiment. The three treatments are applied to each subject, one per month.

This diagram shows the main features of a repeated measures design, which are

1. Each subject receives all treatments.
2. The treatments are applied through time (or space). When the treatments are applied in the same order across all subjects, it is impossible to separate treatment effects from sequence effects. Some processes that can cause *sequence effects* are learning, practice, or fatigue—any pattern in the responses across time that occurs without the treatment. If you think the possibility for sequence effects exists, you must make sure that the effects of prior treatments have been washed out before applying the next treatment.

3. Unlike other designs, the repeated measures design has two experimental units: *between* and *within*. In this example, the first (between) experimental unit is a subject. Subject-to-subject variability is used to test the between factor (groups). The second (within) experimental unit is the time period. In the above example, the month-to-month variability within a subject is used to test the treatment. The important point to realize is that the repeated measures design has two error components, the between and the within.

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 independent, assumption 3 must be modified for responses within a subject. 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.
5. When using an *F*-test, the within-subject covariance matrices are assumed to be *circular*. One way of defining circularity is that the variances of differences between any two measurements within a subject are constant for all measurements. Since responses that are close together in time (or space) often have a higher correlation than those that are far apart, it is common for this assumption to be violated. This assumption is not necessary for the validity of the three multivariate tests: Wilks' lambda, Pillai-Bartlett trace, or Hotelling-Lawley trace.

Advantages of Within-Subjects Designs

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. By reducing the underlying variability, the same power can be achieved with a smaller number of subjects.

Disadvantages of Within-Subjects Designs

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 should try to balance the number of subjects receiving each treatment first.
2. *Carryover effect*. In many drug studies, it is important to wash out the influence of 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.
3. *Statistical analysis.* The statistical model is more restrictive than in a regular factorial design since the individual responses must have certain mathematical properties.

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 so you can make sure that the design is appropriate, rather than learning of them later after the research has been conducted.

Technical Details

General Linear Multivariate Model

This section provides the technical details of the repeated measures designs that can be analyzed by **PASS**. Earlier editions of **PASS** calculated approximate power using the formulas in Muller, LaVange, Ramey, and Ramey (1992). The univariate tests have been updated to use the more accurate formulas found in Muller, Edwards, Simpson, and Taylor (2007).

For *N* subjects, the usual general linear multivariate model is

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

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

$$row_k(R) \sim N_p(0, \Sigma)$$

Note that *p* is the product of the number of levels of each of the within-subject factors, *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

$$H_0\colon \underset{a\times b}{\Theta} = \Theta_0$$

$$\underset{a\times q\times p\times b}{CMD} = \Theta$$

where *C* and *D* are orthonormal contrast matrices and Θ_0 is a matrix of hypothesized values, usually zeros. Note that *C* defines contrasts among the between-subject factor levels and *D* defines contrast among the within-subject factor levels.

Tests of the various main effects and interactions may be constructed with suitable choices for *C* and *D*. These tests are based on

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

$$\widehat{\Theta} = C\widehat{M}D$$

$$\underset{b\times b}{H} = \left(\widehat{\Theta} - \Theta_0\right)'[C(X'X)^{-}C']^{-1}\left(\widehat{\Theta} - \Theta_0\right)$$

$$\underset{b\times b}{E} = D'\widehat{\Sigma}D \cdot (N - r)$$

$$\underset{b\times b}{T} = H + E$$

where *r* is the rank of *X*.

Notation

The formulas below use the following notation and terms:

$Y = XM + R$: general linear multivariate model

N: number of subjects

p: product of the number of levels of all within subject factors

q: number of design variables

Y: $N \times p$ matrix of responses

X: $N \times q$ design matrix

$r = \text{rank}(X)$

$v_e = N - r$

M: $q \times p$ matrix of regression parameters

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

R: $N \times p$ matrix of residuals

$\hat{R} = Y - X\hat{M}$

C: $a \times q$ fixed, known, between subject contrast matrix

D: $p \times b$ fixed, known, within subject contrast matrix

Θ: $a \times b$ matrix of secondary regression parameters

Θ_0: values of Θ under H_0

Σ: Error covariance matrix

$$\hat{\Sigma}_{(p \times p)} = \frac{\hat{R}'\hat{R}}{v_e}$$

$\Sigma_* = D'\Sigma D$

$\hat{\Sigma}_* = D'\hat{\Sigma} D$

$\Theta = CMD$

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

$H = \left(\hat{\Theta} - \Theta_0\right)'[C(X'X)^{-}C']^{-1}\left(\hat{\Theta} - \Theta_0\right)$: $b \times b$ hypothesis sum of squares matrix

$E = S_E = v_e \hat{\Sigma}_* = v_e D' \hat{\Sigma} D$: $b \times b$ error sum of squares matrix

$T = H + \mathrm{E}$: $b \times b$ total sum of squares matrix

$m = \min(a, b)$

$$g_1(v_e, a, b) = \frac{[v_e^2 - v_e(2b+3) + b(b+3)](ab+2)}{v_e(a+b+1) - (a+2b+b^2-1)} + 4$$

$$g_2(v_e, a, b) = \frac{v_e + m - b}{v_e + a}\left[\frac{m(v_e + m - b)(v_e + a + 2)(v_e + a - 1)}{v_e(v_e + a - b)} - 2\right]$$

$$g_3(v_e, a, b) = \begin{cases} g_4(v_e, a, b) & \text{if } a^2b^2 \le 4 \\ g_4(v_e, a, b)\sqrt{\dfrac{a^2b^2 - 4}{a^2 + b^2 - 5}} & \text{if } a^2b^2 > 4 \end{cases}$$

$$g_4(v_e, a, b) = [v_e - (b - a + 1)/2] - (ab - 2)/2$$

$$\varepsilon = \frac{\mathrm{tr}^2(\Sigma_*)}{b\,\mathrm{tr}(\Sigma_*^2)} = \frac{\left(\sum_{k=1}^{b} \lambda_k\right)^2}{b \sum_{k=1}^{b} \lambda_k^2}$$

$$\Sigma_* = D'\Sigma D = \Upsilon \mathrm{diag}(\lambda) \Upsilon'$$

$\lambda = [\lambda_1 \ \lambda_2 \ldots \lambda_b]' =$ vector of eigenvalues of Σ_*

$\Upsilon = [v_1 \ v_2 \ldots v_b] =$ matrix of eigenvectors of Σ_*

$$\Omega = \Delta \Sigma_*^{-1}$$

$$\Omega_t = \mathrm{H}_t \mathrm{diag}(\lambda)^{-1}$$

$$\omega_k = \upsilon_k' \mathrm{H} \upsilon_k / \lambda_k$$

$$S_{t1} = \sum_{k=1}^{b} \lambda_k$$

$$S_{t2} = \sum_{k=1}^{b} \lambda_k \, \omega_{*k}$$

$$S_{t3} = \sum_{k=1}^{b} \lambda_k^2$$

$$S_{t4} = \sum_{k=1}^{b} \lambda_k^2 \, \omega_{*k}$$

$$\lambda_{*1} = (aS_{t3} + 2S_{t4})/(aS_{t1} + 2S_{t2})$$

$$\lambda_{*2} = S_{t3}/S_{t1}$$

$$v_{*1} = aS_{t1}/\lambda_{*1}$$

$$v_{*2} = \frac{v_e S_{t1}^2}{S_{t3}} = v_e b\varepsilon$$

$$\omega_u = S_{t2}/\lambda_{*1}$$

$$\mathrm{E}(\hat{\varepsilon}) \approx b^{-1}\mathrm{E}(t_1)/\mathrm{E}(t_2)$$

$$\mathrm{E}(\tilde{\varepsilon}) \approx b^{-1}[\mathrm{N}\ \mathrm{E}(t_1) - 2\mathrm{E}(t_2)]/[v_e\ \mathrm{E}(t_2) - \mathrm{E}(t_1)]$$

$$\mathrm{E}(t_1) = 2v_e S_{t3} + v_e^2 S_{t1}^2$$

$$\mathrm{E}(t_2) = v_e(v_e + 2)S_{t3} + 2v_e \sum_{k_1=2}^{b} \sum_{k_2=1}^{k_1-1} \lambda_{k_1}\lambda_{k_2}$$

Table 1. Summary of the calculation of each of the statistical tests analyzed in this procedure.

Test Name	Test Type	Statistic	df1	df2
Uncorrected F	Univariate	$\mathrm{tr}(H)/\mathrm{tr}(E)$	ab	v_eb
Geisser-Greenhouse	Univariate	$\mathrm{tr}(H)/\mathrm{tr}(E)$	ab$\tilde{\varepsilon}$	v_eb$\hat{\varepsilon}$
Huynh-Feldt	Univariate	$\mathrm{tr}(H)/\mathrm{tr}(E)$	ab$\tilde{\varepsilon}$	v_eb$\tilde{\varepsilon}$
Box	Univariate	$\mathrm{tr}(H)/\mathrm{tr}(E)$	a	v_e
Hotelling-Lawley	Multivariate	$\mathrm{tr}(HE^{-1})$	ab	$g_1(v_e, a, b)$
Pillai-Bartlett	Multivariate	$\mathrm{tr}(H(H+E)^{-1})$	$\mathrm{ab}\dfrac{g_2(v_e, a, b)}{\mathrm{m}(v_e + \mathrm{m} - b)}$	$g_2(v_e, a, b)$
Wilks' Lambda	Multivariate	$\lvert E(H+E)^{-1}\rvert$	ab	$g_3(v_e, a, b)$

Uncorrected F-Test

Assuming that Σ has compound symmetry, a size α test of $\mathrm{H}_0\colon \Theta = \Theta_0$ is given by the test statistic f_u.

$$f_u = \frac{\mathrm{tr}(H)/a}{\mathrm{tr}(E)}$$

The critical value of this test statistic, f_0(F), is based on the F distribution as follows.

$$f_0(\mathrm{F}) = F_F^{-1}(1-\alpha; ab, bv_e)$$

Using this critical value, the power is calculated as

$$\Pr\{f_u \le f_0(\mathrm{F})\} = F_F(f_0(\mathrm{F}); ab, bv_e, abf_u)$$

Geisser-Greenhouse F-Test

The assumption that Σ has compound symmetry is usually not viable. Box (1954a,b) suggested that adjusting the degrees of freedom of the above *F*-ratio could compensate for the lack of compound symmetry in Σ. His adjustment has become known as the Geisser-Greenhouse adjustment.

Assuming that Σ has compound symmetry, a size α test of $\mathrm{H}_0\colon \Theta = \Theta_0$ is given by the test statistic f_u.

$$f_u = \frac{\mathrm{tr}(S_H)/a}{\mathrm{tr}(S_E)}$$

The expected critical value of this test statistic, f_0(GG), is approximated using the F distribution as follows.

$$f_0(\mathrm{GG}) = F_F^{-1}(1-\alpha; ab\mathrm{E}(\hat{\varepsilon}), bv_e\mathrm{E}(\hat{\varepsilon}))$$

Using this critical value, the power is calculated as

$$\Pr\{f_u \le f_0(\mathrm{GG})\} \approx F_F\left(f_0(\mathrm{GG})\frac{\lambda_{*2}}{\lambda_{*1}}\frac{ab}{v_{*1}}\frac{v_{*2}}{bv_e}; v_{*1}, v_{*2}, \omega_u\right)$$

Note that the Geisser-Greenhouse adjustment is only needed for testing main effects and interactions involving within-subject factors. Main effects and interactions that involve only between-subject factors need no such adjustment.

Huynh-Feldt Test

The assumption that Σ has compound symmetry is usually not viable. Box (1954a,b) suggested that adjusting the degrees of freedom of the above *F*-ratio could compensate for the lack of compound symmetry in Σ. His adjustment has become known as the Geisser-Greenhouse adjustment. This adjustment was further refined by Huynh and Feldt (1970), and it is popular still today.

Even if Σ does not exhibit compound symmetry, an approximate size α test of $\mathrm{H}_0\colon \Theta = \Theta_0$ is given by the test statistic f_u.

$$f_u = \frac{\mathrm{tr}(S_H)/a}{\mathrm{tr}(S_E)}$$

The expected critical value of this test statistic with the Huynh-Feldt adjustment, f_0(HF), is approximated using the F distribution as follows.

$$f_0(\mathrm{HF}) = F_F^{-1}(1-\alpha; ab\mathrm{E}(\tilde{\varepsilon}), bv_e\mathrm{E}(\tilde{\varepsilon}))$$

Using this critical value, the power is calculated as

$$\Pr\{f_u \le f_0(\mathrm{HF})\} \approx F_F\left(f_0(\mathrm{HF})\frac{\lambda_{*2}}{\lambda_{*1}}\frac{ab}{v_{*1}}\frac{v_{*2}}{bv_e}; v_{*1}, v_{*2}, \omega_u\right)$$

Note that the Huynt-Feldt adjustment is only needed for testing main effects and interactions involving within-subject factors. Main effects and interactions that involve only between-subject factors need no such adjustment.

Box's Conservative Test

The assumption that Σ has compound symmetry is usually not viable. Box suggested that adjusting the degrees of freedom of the above *F*-ratio could compensate for the lack of compound symmetry in Σ. His adjustment has become known as the Box's conservative adjustment.

Even if Σ does not exhibit compound symmetry, an approximate size α test of $\mathrm{H}_0\colon \Theta = \Theta_0$ is given by the test statistic f_u.

$$f_u = \frac{\mathrm{tr}(S_H)/a}{\mathrm{tr}(S_E)}$$

The expected critical value of this test statistic with Box's conservative adjustment, f_0(HF), is approximated using the F distribution as follows.

$$f_0(\mathrm{B}) = F_F^{-1}(1-\alpha; a, v_e)$$

Using this critical value, the power is calculated as

$$\Pr\{f_u \le f_0(\mathrm{B})\} \approx F_F\left(f_0(\mathrm{B})\frac{\lambda_{*2}}{\lambda_{*1}}\frac{ab}{v_{*1}}\frac{v_{*2}}{bv_e}; v_{*1}, v_{*2}, \omega_u\right)$$

Note that Box's conservative adjustment is only needed for testing main effects and interactions involving within-subject factors. Main effects and interactions that involve only between-subject factors need no such adjustment.

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

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

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

$$df1 = ab$$

$$df2 = g[(N-r) - (b-a+1) \;/\; 2] - (ab-2) \;/\; 2$$

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

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 non-central *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

$$\lambda = df_1 F_{df1,df2}$$

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

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

$$df1 = ab$$

$$df2 = s[(N-r) - b + s]$$

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

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

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

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

$$df1 = ab$$

$$df2 = s[(N-r) - b - 1] + 2$$

The M (Means) Matrix

In the general linear multivariate model presented above, *M* represents a matrix of regression coefficients. Since you must provide the elements of *M*, we will discuss its meaning in more detail. 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 between-subject categories and the columns of *M* represent the within-group categories.

The *q* rows of *M* represent the *q* groups into which the subjects can be classified. For example, if a design includes three between-subject factors with 2, 3, and 4 categories, the matrix *M* would have 2 *x* 3 *x* 4 = 24 rows. That is, *q* = 24. Similarly, if a design has three within-subject factors with 3, 3, and 3 categories, the matrix *M* would have 3 *x* 3 *x* 3 = 27 columns. That is, *p* = 27.

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 is measured four times. 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 measurement of subjects in the first group. To calculate the power of this design, you would need to specify appropriate values of all twelve means.

As a second example, consider a design with three between-subject factors and three within-subject factors, all of which have two categories. The *M* matrix for this design would be as follows.

$$M = \left[\begin{array}{ccccccccccc}
 & & W1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\
 & & W2 & 1 & 1 & 2 & 2 & 1 & 1 & 2 & 2 \\
 & & W3 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\
B1 & B2 & B3 & & & & & & & & \\
1 & 1 & 1 & \mu_{111111} & \mu_{111112} & \mu_{111121} & \mu_{111122} & \mu_{111211} & \mu_{111212} & \mu_{111221} & \mu_{111222} \\
1 & 1 & 2 & \mu_{112111} & \mu_{112112} & \mu_{112121} & \mu_{112122} & \mu_{112211} & \mu_{112212} & \mu_{112221} & \mu_{112222} \\
1 & 2 & 1 & \mu_{121111} & \mu_{121112} & \mu_{121121} & \mu_{121122} & \mu_{121211} & \mu_{121212} & \mu_{121221} & \mu_{121222} \\
1 & 2 & 2 & \mu_{122111} & \mu_{122112} & \mu_{122121} & \mu_{122122} & \mu_{122211} & \mu_{122212} & \mu_{122221} & \mu_{122222} \\
2 & 1 & 1 & \mu_{211111} & \mu_{211112} & \mu_{211121} & \mu_{211122} & \mu_{211211} & \mu_{211212} & \mu_{211221} & \mu_{211222} \\
2 & 1 & 2 & \mu_{212111} & \mu_{212112} & \mu_{212121} & \mu_{212122} & \mu_{212211} & \mu_{212212} & \mu_{212221} & \mu_{212222} \\
2 & 2 & 1 & \mu_{221111} & \mu_{221112} & \mu_{221121} & \mu_{221122} & \mu_{221211} & \mu_{221212} & \mu_{221221} & \mu_{221222} \\
2 & 2 & 2 & \mu_{222111} & \mu_{222112} & \mu_{222121} & \mu_{222122} & \mu_{222211} & \mu_{222212} & \mu_{222221} & \mu_{222222}
\end{array}\right]$$

The subscripts for each mean follow the pattern $\mu_{B1\,B2\,B3\,W1\,W2\,W3}$. The first three subscripts indicate the between-subject categories, and the second three subscripts indicate the within-subject categories. Notice that the first three subscripts are constant in each row and the second three subscripts are constant in each column.

Specifying the M Matrix

When computing the power in a repeated measures analysis of variance, the specification of the *M* matrix is one of your main tasks. The program cannot do this for you. The calculated power is directly related to your choice. So, your choice for the elements of *M* must be selected carefully and thoughtfully. When authorization and approval from a government organization is sought, you should be prepared to defend your choice of *M*. In this section, we will explain how you can specify *M*.

Before we begin, it is important that you have in mind exactly what *M* is. *M* is a table of means that represent the size of the differences among the means that you want the study or experiment to detect. That is, *M* gives the means under the alternative hypothesis. Under the null hypothesis, these means are assumed to be equal. Because of the complexity of the repeated measures design, it is often difficult to choose reasonable values, so **PASS** will help you. But it is important to remember that you are responsible for these values and that the sample sizes calculated are based on them.

One way to specify the *M* matrix is to do so directly into the spreadsheet. You might do this if you are calculating the 'retrospective' power of a study that has already been completed, or if it is simply easier to write the matrix directly. Usually, however, you will specify the *M* matrix in portions.

We will begin our discussion of specifying the *M* matrix with an example. Consider a study of two groups of subjects. Each subject was tested, then a treatment was administered, then the subject was tested again at the ten-minute mark, and then tested a third time after sixty minutes. The researchers wanted the sample size to be large enough to detect the following pattern in the means.

Table of Hypothesized Means				
	Time Period			
Group	T0	T10	T60	Average
A	100	130	100	110
B	120	180	120	140
Average	110	155	110	125

To understand how they derived this table, we will perform some basic arithmetic on it.

Step 1 – Remove the Overall Mean

Subtract 125, the overall mean, from each of the individual means.

Table of Hypothesized Means Adjusted for Overall Mean				
	Time Period			
Group	T0	T10	T60	Average
A	-25	5	-25	-15
B	-5	55	-5	15
Average	-15	30	-15	125

Step 2 – Remove the Group Effect

Subtract -15 from the first row and 15 from the second row.

Table of Hypothesized Means Adjusted for Group				
	Time Period			
Group	T0	T10	T60	Total
A	-10	20	-10	-15
B	-20	40	-20	15
Total	-15	30	-15	125

Step 3 – Remove the Time Effect

Subtract -15 from the first column, 30 from the second column, and -15 from the third column.

Table of Hypothesized Means Adjusted for Group and Time					
	Time Period				
Group	T0	T10	T60	Effect	Effect + Overall
A	5	-10	5	-15	110
B	-5	10	-5	15	140
Effect	-15	30	-15		
Effect + Overall	110	155	110		125

This table, called an effects table, lets us see the individual effect of each component of the model. For example, we can see that the hypothesized pattern across time is that T10 is 45 units higher than either endpoint. Similarly, we note that the hypothesized pattern for the two groups is that Group B is 30 units larger than Group A.

Understanding the interaction is more difficult. One interpretation focuses on T10. We note that in Group A the response for T10 is 10 less than expected while in Group B the response for T10 is 10 more than expected.

The C Matrix for Between-Subject Contrasts

The *C* matrix is comprised of contrasts that are applied to the rows of *M*. That is, these are between-group contrasts. 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 between-subject term in the model. For example, in the six factor example above, the *C* matrix that will be generated for testing the between-subject factor B1 is

$$C_{B1} = \begin{bmatrix} \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \end{bmatrix}$$

Note that the divisor $\sqrt{8}$ is used so that the total of the squared elements is one. This is required so that the contrast matrix is *orthonormal*.

When creating a test for B1, the matrix *D* is created to average across all within-subject categories.

$$D_{B1} = \begin{bmatrix} \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} \end{bmatrix}$$

Generating the C Matrix when There are Multiple Between Factors

Generating the *C* matrix when there is more than one between factor is more difficult. We like 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

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

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

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

Step 3. Create the appropriate contrast matrix using a direct (Kronecker) product of either the $\ddot{C}_{Bi}$ 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 *B*2

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

Expanding the direct product results in

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

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-2}{\sqrt{12}} & \frac{-2}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ 0 & 0 & \frac{-1}{\sqrt{4}} & \frac{-1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix} \otimes \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{-2}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} & \frac{1}{\sqrt{48}} \\ 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & 0 & 0 & \frac{-1}{\sqrt{16}} & \frac{-1}{\sqrt{16}} & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{16}} \end{bmatrix}$$

Similarly, the *C* matrix suitable for testing interaction *B*2*B*3 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.

The D Matrix for Within-Subject Contrasts

The *D* matrix is comprised of contrasts that are applied to the columns of *M*. That is, these are within-group contrasts. You do not have to specify these contrasts either. They will be generated for you. Specification of the *D* matrix is similar to the specification of the *C* matrix, except that now the matrices are all transposed.

Interactions of Between-Subject and Within-Subject Factors

Interactions that include both between-subject factors and within-subject factors require that between-subject portion be specified by the *C* matrix and the within-subject portion be specified with the *D* matrix.

Covariance Matrix Assumptions

The following assumptions are made when using the *F*-test. These assumptions are not needed when using one of the three multivariate tests.

In order to use the *F* ratio to test hypotheses, certain assumptions are made about the distribution of the residuals e_{ijk}. Specifically, it is assumed that the residuals for each subject, $e_{ij1}, e_{ij2}, \cdots, e_{ijT}$, are distributed as a multivariate normal with means equal to zero and covariance matrix Σ_{ij}. Two additional assumptions are made about these covariance matrices. First, they are assumed to be equal for all subjects. That is, it is assumed that $\Sigma_{11} = \Sigma_{12} = \cdots = \Sigma_{Gn} = \Sigma$. Second, the covariance matrix is assumed to have a particular form called *circularity*. A covariance matrix is *circular* if there exists a matrix *A* such that

$$\Sigma = A + A' + \lambda I_T$$

where I_T is the identity matrix of order *T* and λ is a constant.

This property may also be defined as

$$\sigma_{ii} + \sigma_{jj} - 2\sigma_{ij} = 2\lambda$$

One type of matrix that is circular is one that 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

$$\Sigma = \begin{bmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \cdots & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & \cdots & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & \cdots & \rho\sigma^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & \cdots & \sigma^2 \end{bmatrix}$$

or, with actual numbers,

$$\begin{bmatrix} 9 & 2 & 2 & 2 \\ 2 & 9 & 2 & 2 \\ 2 & 2 & 9 & 2 \\ 2 & 2 & 2 & 9 \end{bmatrix}$$

An example of a matrix which does not have compound symmetry but is still circular is

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 4 & 5 \\ 3 & 6 & 5 & 6 \\ 4 & 5 & 8 & 7 \\ 5 & 6 & 7 & 10 \end{bmatrix}$$

Needless to say, the need to have the covariance matrix circular is a very restrictive assumption.

Between-Subject Standard Deviation

The subject-to-subject variability is represented by $\sigma^2_{Between}$. In a repeated measures AOV table, this quantity is estimated by the between subjects mean square (*MSB*). This quantity is calculated from Σ using the formula

$$\sigma^2_{Between} = \frac{\sum_{i=1}^{T}\sum_{j=1}^{T}\sigma_{ij}}{T}$$

$$= \frac{\sum_{i=1}^{T}\sum_{j=1}^{T}\sigma_{ii}\,\sigma_{jj}\rho_{ij}}{T}$$

When Σ has compound symmetry, which requires all $\sigma_{ii} = \sigma$ and all $\rho_{ij} = \rho$, the above formula reduces to

$$\sigma^2_{Between} = \sigma^2(1 + (T-1)\rho)$$

Note that *F*-tests of between factors and their interactions do not require the circularity assumption so the Geisser-Greenhouse correction is not applied to these tests.

Within-Subject Standard Deviation

The within-subject variability is represented by σ^2_{Within}. In a repeated measures AOV table, this quantity is estimated by the within-subjects mean square (*MSW*). This quantity is calculated from Σ using the formula

$$\sigma^2_{Within} = \frac{\sum_{i=1}^{T} \sigma_{ii}}{T} - \frac{2\sum_{i=1}^{T}\sum_{j=i+1}^{T} \sigma_{ij}}{T(T-1)}$$

$$= \frac{\sum_{i=1}^{T} \sigma_{ii}}{T} - \frac{2\sum_{i=1}^{T}\sum_{j=i+1}^{T} \rho_{ij}\sqrt{\sigma_{ii}\sigma_{jj}}}{T(T-1)}$$

When Σ has compound symmetry, which requires all $\sigma_{ii} = \sigma$ and all $\rho_{ij} = \rho$, the above formula reduces to

$$\sigma^2_{Within} = \sigma^2(1-\rho)$$

Estimating Sigma and Rho from Existing Data

Using the above results for existing data, approximate values for σ and ρ may be estimated from a previous analysis of variance table that provides estimates of MSB and MSW. Solving the above equations for σ and ρ yields

$$\rho = \frac{\sigma^2_{Between} - \sigma^2_{Within}}{\sigma^2_{Between} + (T-1)\sigma^2_{Within}}$$

$$\sigma^2 = \frac{\sigma^2_{Within}}{1-\rho}$$

Substituting MSB for $\sigma^2_{Between}$ and MSW for σ^2_{Within} yields the estimates

$$\hat{\rho} = \frac{MSB - MSW}{MSB + (T-1)MSW}$$

$$\hat{\sigma}^2 = \frac{MSW}{1-\hat{\rho}}$$

Note that these estimators assume that the design meets the circularity assumption, which is usually not the case. However, they provide crude estimates that can be used in planning.

Notes for Some Procedure Options

Columns Containing the Means

This option appears if Mean Input Type is set to 'Means, Interaction Effects.' Use this option to the specify spreadsheet columns containing a hypothesized matrix of mean from which the value of σm 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 between factors are represented across the columns and the within factors 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 between factors. If no between factors are specified, the number of groups is one.

The number of rows with data in these columns must equal the number of times a subject is measured. Thus, the number of rows is equal to the product of the number of levels of the within factors. If no within factors are specified, the number of rows is one.

For example, suppose you are designing an experiment that is to have two between factors (A & B) and two within factors (D & E), each with two levels. The four columns of the spreadsheet would be

A1B1 A1B2 A2B1 A2B2

The rows of the spreadsheet would represent

D1E1

D1E2

D2E1

D2E2

Example

To see how this option works, consider the following table of hypothesized means for an experiment with one between factor (A) having two groups and one within factor (B) having three time periods. Suppose the values in columns C1 and C2 of the spreadsheet are

C1	C2
2.0	4.0
4.0	6.0
6.0	11.0

The following table of effects results from forming and subtracting the row and column means.

	C1	C2	Means	Effects
Row1	0.5	-0.5	3.0	-2.5
Row2	0.5	-0.5	5.0	-0.5
Row3	-1.0	1.0	8.5	3.0
	---	---	---	
Means	4.0	7.0	5.5	
Effects	-1.5	1.5		

The standard deviation of the A effects is calculated as

$$\sigma_A = \sqrt{\frac{(-1.5)^2 + (1.5)^2}{2}}$$

$$= \sqrt{2.25}$$

$$= 1.5$$

The standard deviation of the B effects is calculated as

$$\sigma_B = \sqrt{\frac{(-2.5)^2 + (-0.5)^2 + (3.0)^2}{3}}$$

$$= \sqrt{\frac{15.5}{3}}$$

$$= 2.27$$

The standard deviation of the interaction effects is found to be

$$\sigma_{AB} = \sqrt{\frac{(0.5)^2 + (0.5)^2 + (-1.0)^2 + (-0.5)^2 + (-0.5)^2 + (1.0)^2}{6}}$$

$$= \sqrt{\frac{3.0}{6}}$$

$$= 0.71$$

These three standard deviations are used to represent the effect sizes of the corresponding terms.

Autocorrelation Structure

Specify the autocorrelation structure of the matrix associated with each factor. The number of diagonal elements in the matrix is equal to the number of levels in the factor. The final autocorrelation matrix is the Kronecker product of these individual factor autocorrelation matrices. This method was presented in Naik and Rao (2001).

For example, suppose an experiment is being designed with two within factors: 1) three equal-space time points and 2) two locations in the brain. Suppose that an AR(1) pattern is assumed for the time factor with autocorrelation γ = 0.6 and a constant pattern is assumed for the location factor with autocorrelation θ = 0.1. The individual factor autocorrelation matrices would be

$$A = \begin{bmatrix} 1 & \gamma & \gamma^2 \\ \gamma & 1 & \gamma \\ \gamma^2 & \gamma & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.6 & 0.36 \\ 0.6 & 1 & 0.6 \\ 0.36 & 0.6 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$$

And the final autocorrelation structure would be

$$A \otimes B = \begin{bmatrix} 1 & \theta & \gamma & \gamma\theta & \gamma^2 & \theta\gamma^2 \\ \theta & 1 & \gamma\theta & \gamma & \theta\gamma^2 & \gamma^2 \\ \gamma & \gamma\theta & 1 & \theta & \gamma & \gamma\theta \\ \gamma\theta & \gamma & \theta & 1 & \gamma\theta & \gamma \\ \gamma^2 & \theta\gamma^2 & \gamma & \gamma\theta & 1 & \theta \\ \theta\gamma^2 & \gamma^2 & \gamma\theta & \gamma & \theta & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.100 & 0.600 & 0.060 & 0.360 & 0.036 \\ 0.100 & 1 & 0.060 & 0.600 & 0.036 & 0.360 \\ 0.600 & 0.060 & 1 & 0.100 & 0.600 & 0.060 \\ 0.060 & 0.600 & 0.100 & 1 & 0.060 & 0.600 \\ 0.360 & 0.036 & 0.600 & 0.060 & 1 & 0.100 \\ 0.036 & 0.360 & 0.060 & 0.600 & 0.100 & 1 \end{bmatrix}$$

We will now present the various options available for quickly specifying the autocorrelation structure of an individual factor.

Constant

A single value of ρ is used as the autocorrelation for all off-diagonal elements of the matrix. This matrix pattern is called compound symmetry.

The matrix appears as follows:

$$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

AR(1)

A single value of ρ is used to generate a first order autocorrelation pattern. This pattern reduces the autocorrelation at each successive step by multiplying the value at the last step by ρ. The times (or locations) are assumed to be equi-spaced.

The matrix appears as follows:

$$\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

LEAR

A single value of ρ and a dampening value δ are used to generate the autocorrelations using the LEAR (linear exponent autoregressive) correlation structure proposed by Simpson, Edwards, Muller, Sen, and Styner (2010). The times (or locations) need to be specified. The formula for this structure is

$$\rho^A, \text{where } A = d_{min} + \delta\left(\frac{d_{jk} - d_{min}}{d_{max} - d_{min}}\right)$$

where $d_{jk} = |t_j - t_k|$, t_j and t_k are any two measurement points (times or locations), d_{min} is the minimum d_{jk}, d_{max} is the maximum d_{jk}, and δ is a dampening constant.

The *t's* often are entered as increasing integers, such as 1, 2, 3, and so on. But this is not necessary. The only required characteristic is that they be strictly increasing. *The* authors recommend that the *t's* be scaled so that $d_{min} = 1$.

Banded

A list of correlation values ρ_1, ρ_2, ... is used to create a banded correlation matrix. If note enough values are entered, the last entered value is carried forward.

A banded correlation matrix for a factor with six levels looks as follows:

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_3 & \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_4 & \rho_3 & \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_5 & \rho_4 & \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$$

Unstructured

An unstructured (no special pattern) correlation matrix is loaded from the specified columns of the spreadsheet. The matrix may be upper-, or lower-, triangular.

The resulting matrix looks like

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 \\ \rho_1 & 1 & \rho_5 & \rho_6 & \rho_7 & \rho_8 \\ \rho_2 & \rho_5 & 1 & \rho_9 & \rho_{10} & \rho_{11} \\ \rho_3 & \rho_6 & \rho_9 & 1 & \rho_{12} & \rho_{13} \\ \rho_4 & \rho_7 & \rho_{10} & \rho_{12} & 1 & \rho_{14} \\ \rho_5 & \rho_8 & \rho_{11} & \rho_{13} & \rho_{14} & 1 \end{bmatrix}$$

The Resulting Variance-Covariance Matrix

Finally, the σ's and the autocorrelation matrix are combined to form the variance-covariate matrix. In general, it will appear as follows.

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_6 \end{bmatrix} \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_3 & \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_4 & \rho_3 & \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_5 & \rho_4 & \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_6 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \rho_3\sigma_1\sigma_4 & \rho_4\sigma_1\sigma_5 & \rho_5\sigma_1\sigma_6 \\ \rho_1\sigma_1\sigma_2 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 & \rho_3\sigma_2\sigma_5 & \rho_4\sigma_2\sigma_6 \\ \rho_2\sigma_1\sigma_3 & \rho_1\sigma_2\sigma_3 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 & \rho_2\sigma_3\sigma_5 & \rho_3\sigma_3\sigma_6 \\ \rho_3\sigma_1\sigma_4 & \rho_2\sigma_2\sigma_4 & \rho_1\sigma_3\sigma_4 & \sigma_4^2 & \rho_1\sigma_4\sigma_5 & \rho_2\sigma_4\sigma_6 \\ \rho_4\sigma_1\sigma_5 & \rho_3\sigma_2\sigma_5 & \rho_2\sigma_3\sigma_5 & \rho_1\sigma_4\sigma_5 & \sigma_5^2 & \rho_1\sigma_5\sigma_6 \\ \rho_5\sigma_1\sigma_6 & \rho_4\sigma_2\sigma_6 & \rho_3\sigma_3\sigma_6 & \rho_2\sigma_4\sigma_6 & \rho_1\sigma_5\sigma_6 & \sigma_6^2 \end{bmatrix}$$

Example 1 – Determining Sample Size

Researchers are planning a study of the impact of a drug on heart rate. They want to evaluate the differences in heart rate among three age groups: 20-40, 41-60, and over 60. Their experimental protocol calls for a baseline heart rate measurement, followed by administration of a certain level of the drug, followed by three additional measurements 30 minutes apart. They want to be able to detect a 10% difference in heart rate among the age groups. They want to detect 5% difference in heart rate within an individual across time. They decide the experiment should detect interaction effects of the same magnitude as the within factor. They plan to analyze the data using a Geisser-Greenhouse corrected F-test.

Similar studies have found an average heart rate of 93, a standard deviation of 4.0, and an autocorrelation between adjacent measurements on the same individual of 0.7. The researchers assume that first-order autocorrelation adequately represents the autocorrelation pattern.

From a heart rate of 93, a 10% reduction gives 84. They decide on the age-group means of 93, 87, and 84. Similarly, a 5% reduction within a subject would result in a heart rate of 88. They decide on time means of 93, 89, 88, and 91.

How many subjects per age group are needed to achieve 95% power and a 0.05 significance level for all terms?

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For	**Sample Size**
Based on	**Geisser-Greenhouse**
Term	**All**
Geisser-Greenhouse	**Checked**
Minimum Power	**0.95**
Alpha for All Terms	**0.05**
Number of Factors	**1 Between, 1 Within**
Mean Input Type	**Means, Interaction Effects**
Factor B1 Levels	**3**
Factor B1 Means	**List of Means**
Factor B1 List of Means	**93 87 84**
Factor W1 Levels	**4**
Factor W1 Means	**List of Means**
Factor W1 List of Means	**93 89 88 91**
K's (Multipliers)	**1.0**
Group Allocation	**Equal (n1 = n2 = ···)**

Interactions Tab

B1*W1 Interaction Enter Effect As ....**Multiple of Another Term**
B1*W1 Interaction Multiplier..........................**1.0**
B1*W1 Interaction Basis Term......................**W1**

Variances Tab

Input Type..**Non-Constant σ's and ρ's**
Pattern ....**Constant σ**
σ **4**
Factor W1 Autocorrelation Structure.............**AR(1)**
Factor W1 ρ **0.7**

Output

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

Design Report

Design Report

Solve For: Sample Size Based on the Geisser-Greenhouse Corrected F-Test (All Terms)

			Sample Size						
Term*	Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (3)	GG F	0.9793	6	18	1	3.74	3.29	1.136	0.05
W1 (4)	GG F	0.9998	6	18	1	1.92	1.31	1.465	0.05
B1*W1	GG F	0.9969	6	18	1	1.92	1.31	1.465	0.05

* The numbers in parentheses represent the number of levels associated with the factor.

Term The identifying label for the factor or interaction. For factors, the number of levels is also given in parentheses.
Test Identifies the test statistic for which the power is calculated.
GG F Geisser-Greenhouse Corrected F-Test
Power The computed power for the term.
n The number of subjects per group.
N The total number of subjects in the study.
K The means/effects were multiplied by this value.
σm Standard Deviation of Effects. This value represents the magnitude of differences among the means for the term.
σ Standard Deviation. The random variation against which σm is compared in the F-test.
σm / σ Effect Size. An index of the size of the mean differences relative to the standard deviation.
Alpha The probability of rejecting a true null hypothesis.

The required sample size is 6 per group. The *Design Report* gives the power for each term. It is useful when you want to compare the powers of the terms in the design at a specific sample size.

Term Reports

Results for Factor B1 (Levels = 3)

Solve For: Sample Size Based on the Geisser-Greenhouse Corrected F-Test (All Terms)

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.9793	6	18	1	3.74	3.29	1.136	0.05

Test Identifies the test statistic for which the power is calculated.
GG F Geisser-Greenhouse Corrected F-Test
Power The computed power for the term.
n The number of subjects per group.
N The total number of subjects in the study.
K The means/effects were multiplied by this value.
σm Standard Deviation of Effects. This value represents the magnitude of differences among the means for the term.
σ Standard Deviation. The random variation against which σm is compared in the F-test.
σm / σ Effect Size. An index of the size of the mean differences relative to the standard deviation.
Alpha The probability of rejecting a true null hypothesis.

Results for Factor W1 (Levels = 4)

Solve For: Sample Size Based on the Geisser-Greenhouse Corrected F-Test (All Terms)

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.9998	6	18	1	1.92	1.31	1.465	0.05

Test Identifies the test statistic for which the power is calculated.
GG F Geisser-Greenhouse Corrected F-Test
Power The computed power for the term.
n The number of subjects per group.
N The total number of subjects in the study.
K The means/effects were multiplied by this value.
σm Standard Deviation of Effects. This value represents the magnitude of differences among the means for the term.
σ Standard Deviation. The random variation against which σm is compared in the F-test.
σm / σ Effect Size. An index of the size of the mean differences relative to the standard deviation.
Alpha The probability of rejecting a true null hypothesis.

Results for Term B1*W1

Solve For: Sample Size Based on the Geisser-Greenhouse Corrected F-Test (All Terms)

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.9969	6	18	1	1.92	1.31	1.465	0.05

Test Identifies the test statistic for which the power is calculated.
GG F Geisser-Greenhouse Corrected F-Test
Power The computed power for the term.
n The number of subjects per group.
N The total number of subjects in the study.
K The means/effects were multiplied by this value.
σm Standard Deviation of Effects. This value represents the magnitude of differences among the means for the term.
σ Standard Deviation. The random variation against which σm is compared in the F-test.
σm / σ Effect Size. An index of the size of the mean differences relative to the standard deviation.
Alpha The probability of rejecting a true null hypothesis.

The *Term Reports* provide a complete report for each term at all sample sizes. They are especially useful when you are only interested in the power of one or two terms.

Geisser-Greenhouse Correction Detail Report

Geisser-Greenhouse Correction Detail Report

Term*	Power	Alpha	Critical F	Lambda	df1	df2	Epsilon	E(Epsilon)	G1
n = 6 N = 18 Means x 1									
B1 (3)	0.9793	0.05	3.68	23.23	2	15	1.00	1.0	0.00
W1 (4)	0.9998	0.05	3.32	38.64	3	45	0.77	0.7	-1.01
B1*W1	0.9969	0.05	2.69	38.64	6	45	0.77	0.7	-1.01

* The numbers in parentheses represent the number of levels associated with the factor.

This report gives the details of the components of the Geisser-Greenhouse correction for each term and sample size. It is useful when you want to compare various aspects of this test.

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

Term

This column contains the identifying label of the term. For factors, the number of levels is also given in parentheses.

Power

This is the computed power for the term.

Alpha

Alpha is the significance level of the test.

Critical F

This is the critical value of the F statistic. An *F* value computed from the data that is larger than this value is statistically significant at the alpha level given.

Lambda

This is the value of the noncentrality parameter λ of the approximate noncentral *F* distribution.

df1 and df2

These are the values of the numerator and denominator degrees of freedom of the approximate *F*-test that is used. These values are useful when comparing various designs. Other things being equal, you would like to have df2 large and df1 small.

Epsilon

The Geisser-Greenhouse epsilon is a measure of how far the covariance matrix departs from the assumption of circularity.

E(Epsilon)

This is the expected value of epsilon. It is a measure of how far the covariance matrix departs from the assumption of circularity.

G1

G1 is part of a correction factor used to convert ε to $E(\hat{\varepsilon})$. It is reported for your convenience.

Summary Statements

Summary Statements

A repeated measures design with 1 between factor (with 3 between factor groups) and 1 within factor (with each subject measured 4 times) will be used to test whether there are differences among the levels of the factors. Each term will be tested using a Geisser-Greenhouse corrected F-test with a Type I error rate (α) of 0.05. The assumed standard deviations, correlations, and variance-covariance matrix are as described in the appropriate sections. Based on the spreads of means described below, to obtain 95% power for all tested terms, the needed number of subjects is 6 per between factor group (for a total of 18). With a sample size of 6 subjects per group, for B1, to detect an effect standard deviation of 3.74 (an effect size of 1.136), the power is 0.9793. For W1, to detect an effect standard deviation of 1.92 (an effect size of 1.465), the power is 0.9998. For B1*W1, to detect an effect standard deviation of 1.92 (an effect size of 1.465), the power is 0.9969.

A summary statement can be generated for each sample size that was entered. This statement gives the results in sentence form. The number of designs reported on textually is controlled by the Summary Statement option on the Reports Tab.

Dropout-Inflated Sample Size Report

Dropout-Inflated Sample Size

Average Group Sample Size n	Group	Dropout Rate	Sample Size ni	Dropout-Inflated Enrollment Sample Size ni'	Expected Number of Dropouts Di
6	1 - 3	20%	6	8	2
	Total		18	24	6

n The average group sample size.
Group Lists the group numbers.
Dropout Rate The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
ni The evaluable sample size for each group at which power is computed. If ni subjects are evaluated out of the ni' subjects that are enrolled in the study, the design will achieve the stated power.
ni' The number of subjects that should be enrolled in each group in order to obtain ni evaluable subjects, based on the assumed dropout rate. After solving for ni, ni' is calculated by inflating ni using the formula ni' = ni / (1 - DR), with ni' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
Di The expected number of dropouts in each group. Di = ni' - ni.

Dropout Summary Statement

Anticipating a 20% dropout rate, group sizes of 8, 8, and 8 subjects should be enrolled to obtain final group sample sizes of 6, 6, and 6 subjects.

This report shows the sample sizes adjusted for dropout. In this example, dropout is assumed to be 20%. You can change the dropout rate on the Reports tab.

Means Matrix

Means Matrix

Name	B1	B2	B3
W1	-10.6	5.2	-2.7
W2	1.5	4.0	2.7
W3	-4.6	4.6	0.0
W4	-4.6	4.6	0.0

This report shows the means matrix that was read in from the spreadsheet or generated from the means and interaction values entered. 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.

Variance-Covariance Matrix

Variance-Covariance Matrix

Name	W1	W2	W3	W4
W1	16.00	11.20	7.84	5.49
W2	11.20	16.00	11.20	7.84
W3	7.84	11.20	16.00	11.20
W4	5.49	7.84	11.20	16.00

This report shows the variance-covariance matrix that was read in from the spreadsheet or generated by the settings of on the Design: Variance tab.

Standard Deviations and Correlation Matrix

Standard Deviations and Correlation Matrix

Name	W1	W2	W3	W4
W1	4.000	0.70	0.49	0.343
W2	0.700	4.00	0.70	0.490
W3	0.490	0.70	4.00	0.700
W4	0.343	0.49	0.70	4.000

SD's on the diagonal. Correlations on the off diagonal(s).

This report shows the standard deviations on the diagonal and the autocorrelations on the off diagonals.

References Section

References

Edwards, L.K. 1993. Applied Analysis of Variance in the Behavior Sciences. Marcel Dekker. New York.
Muller, 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.
Muller, 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.
Muller, K.E., Edwards, L.J., Simpson, S.L., and Taylor, D.J. 2007. 'Statistical tests with accurate size and power for balanced linear mixed models.' Statistics in Medicine, Volume 26, pages 3639-3660.
Simpson, S.L., Edwards, L.J., Muller, K.E., Sen, P.K., and Styner, M.A. 2010. 'A linear exponent AR(1) family of correlation structures.' Statistics in Medicine, Volume 29(17), pages 1825-1838.
Naik, D.N. and Rao, S.S. 2001. 'Analysis of multivariate repeated measures data with a Kronecker product structured covariance matrix.' Journal of Applied Statistics, Volume 28 No. 1, pages 91-105.

This report shows the references for this procedure.

Example 2 – Varying the Difference Between the Means

Continuing with Example 1, the researchers want to evaluate the impact on power of varying the size of the difference among the means for a range of sample sizes from 2 to 8 per groups. The researchers could try calculating various multiples of the means, inputting them, and recording the results. However, this can be accomplished directly by using the *K* option.

Keeping all other settings as in Example 2, the value of *K* is varied from 0.2 to 3.0 in steps of 0.2. We determined these values by experimentation so that a full range of power values are shown on the plots.

In the output to follow, we only display the plots. You may want to display the numeric reports as well, but we do not here in order to save space.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Power**
Geisser-Greenhouse.. **Checked**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type... **Means, Interaction Effects**
Factor B1 Levels.. **3**
Factor B1 Means .... **List of Means**
Factor B1 List of Means................................... **93 87 84**
Factor W1 Levels... **4**
Factor W1 Means .... **List of Means**
Factor W1 List of Means.................................. **93 89 88 91**
K's (Multipliers) .... **0.2 to 3.0 by 0.2**
Group Allocation .... **Equal (n1 = n2 = ⋯ = n)**
n (Size Per Group)... **2 3 4 8**

Interactions Tab

B1*W1 Interaction Enter Effect As .... **Multiple of Another Term**
B1*W1 Interaction Multiplier........................... **1.0**
B1*W1 Interaction Basis Term....................... **W1**

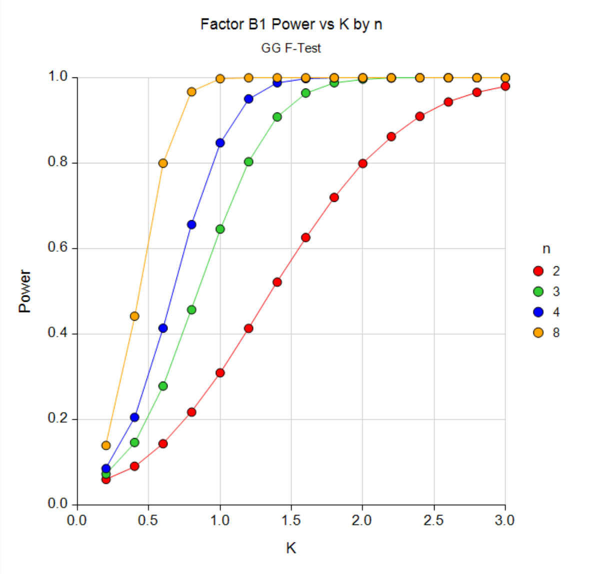
Variances Tab

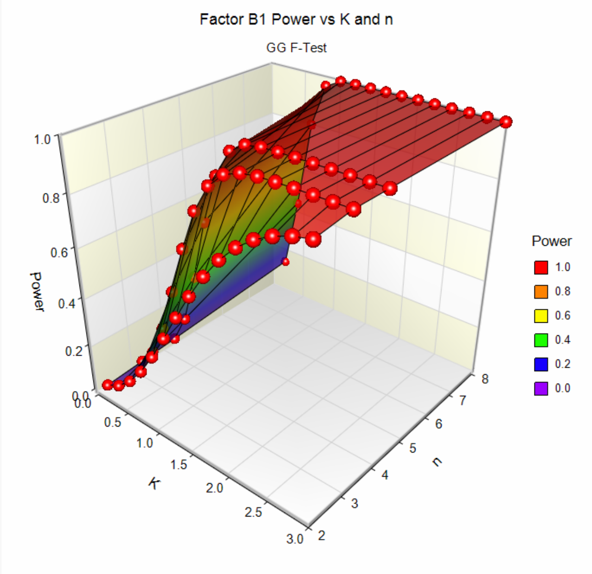
Input Type... **Non-Constant σ's and ρ's**
Pattern .... **Constant σ**
σ **4**
Factor W1 Autocorrelation Structure.............. **AR(1)**
Factor W1 ρ **0.7**

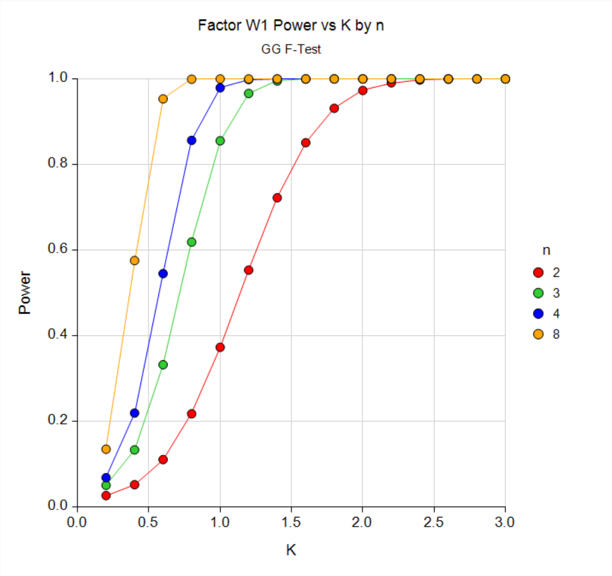
Output

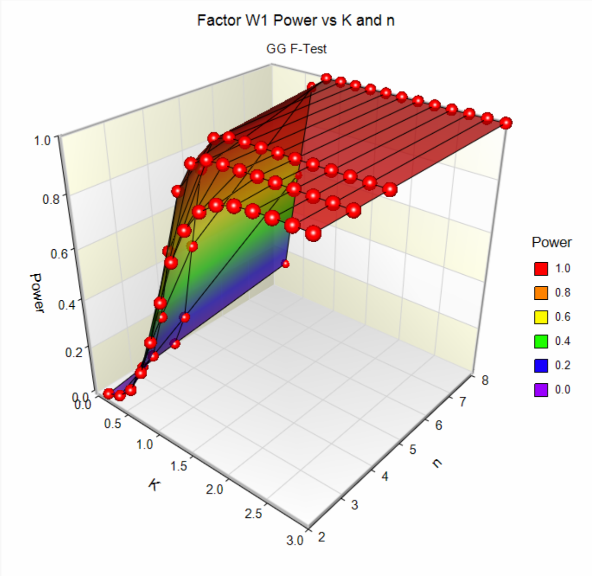
Click the Calculate button to perform the calculations and generate the following output.

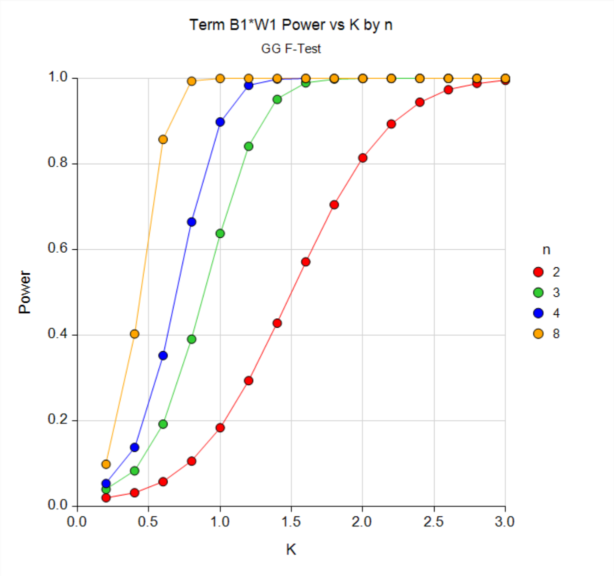
Plots

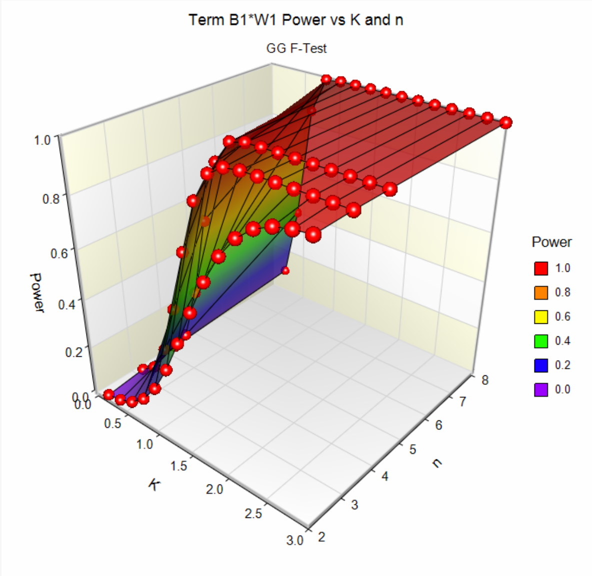












These charts show how the power depends on the relative size of the means as well as the group sample size *n*.

Example 3 – Power After a Study

This example will show how to calculate the power of *F*-tests from data that have already been collected and analyzed using the analysis of variance. The following results were obtained using the analysis of variance procedure in **NCSS**. In this example, Gender is the between factor with two levels and Treatment is the within factor with three levels. The experiment was conducted with two subjects per group, but there is interest in the power for 2, 3, and 4 subjects per group. All tests use a significance level of 0.05.

Analysis of Variance Table

Model Term	DF	Sum of Squares	Mean Square	F-Ratio	P-Value
A: Gender	1	21.33333	21.33333	32.00	0.029857
B(A): Subject	2	1.333333	0.6666667		
C: Treatment	2	5.166667	2.583333	6.20	0.059488
AC	2	5.166667	2.583333	6.20	0.059488
BC(A)	4	1.666667	0.4166667		
S	0				
Total (Adjusted)	11	34.66667			
Total	12				

Means and Standard Errors

Term	Count	Mean	Standard Error
All	12	17.33333	
A: Gender			
Females	6	16	0.3333333
Males	6	18.66667	0.3333333
C: Treatment			
L	4	16.75	0.3227486
M	4	17	0.3227486
H	4	18.25	0.3227486
AC: Gender, Treatment			
Females, L	2	14.5	0.4564355
Females, M	2	16	0.4564355
Females, H	2	17.5	0.4564355
Males, L	2	19	0.4564355
Males, M	2	18	0.4564355
Males, H	2	19	0.4564355

Note that the treatment means (L, M, and H) show an increasing pattern from 16.75 to 18.25, but the hypothesis test of this factor is not statistically significant at the 0.05 level. We will now calculate the power of the three *F*-tests using **PASS**. We will use the regular F-test since that is what was used in the above table.

Using the means from the table, the following means matrix is created.

14.5	19
16	18
17.5	19

From the printout, we note that MSB = 0.6666667 and MSW = 0.4166667. Plugging these values into the estimating equations

$$\hat{\rho} = \frac{MSB - MSW}{MSB + (T-1)MSW}$$

$$\hat{\sigma}^2 = \frac{MSW}{1 - \hat{\rho}}$$

yields

$$\hat{\rho} = \frac{0.6666667 - 0.4166667}{0.6666667 + (3-1)0.4166667} = 0.16666667$$

$$\hat{\sigma}^2 = \frac{0.4166667}{1 - 0.16666667} = 0.5$$

so that

$$\hat{\sigma} = \sqrt{0.5} = 0.70710681$$

With these values calculated, we can setup **PASS** to calculate the power of the three *F*-tests as follows.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Power**
F-Test .... **Checked**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type .... **Means in Spreadsheet**
Factor B1 Levels **2**
Factor W1 Levels **3**
Columns Containing the Means **C1-C2**
K's (Multipliers) **1.0**
Group Allocation .... **Equal (n1 = n2 = ⋯ = n)**
n (Size Per Group) **2 3 4**

Variances Tab

Input Type .... **Constant σ and ρ**
σ (Standard Deviation) **0.70710681**
ρ (Autocorrelation) **0.16667**

Input Spreadsheet Data

Row	C1	C2
1	14.5	19
2	16.0	18
3	17.5	19

Output

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

Design Report

Solve For: Power

			Sample Size						
Term*	Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (2)	F	0.8004	2	4	1	1.33	0.47	2.828	0.05
W1 (3)	F	0.5536	2	4	1	0.66	0.37	1.761	0.05
B1*W1	F	0.5536	2	4	1	0.66	0.37	1.761	0.05
B1 (2)	F	0.9985	3	6	1	1.33	0.47	2.828	0.05
W1 (3)	F	0.8933	3	6	1	0.66	0.37	1.761	0.05
B1*W1	F	0.8933	3	6	1	0.66	0.37	1.761	0.05
B1 (2)	F	1.0000	4	8	1	1.33	0.47	2.828	0.05
W1 (3)	F	0.9801	4	8	1	0.66	0.37	1.761	0.05
B1*W1	F	0.9801	4	8	1	0.66	0.37	1.761	0.05

* The numbers in parentheses represent the number of levels associated with the factor.

You can see that the power of the tests on W1 and B1*W1 is only 0.5536 for an *n* of 2. However, if *n* is 3, a much more reasonable power of 0.8933 is achieved.

Example 4 – Cross-Over Design

A *crossover design* is a special type of repeated measures design in which the treatments are applied to the subjects in different orders. The between-subjects (grouping) factor is defined by the specific sequence in which the treatments are applied. For example, suppose the treatments are represented by B1 and B2. Further suppose that half the subjects receive treatment B1 followed by treatment B2 (sequence B1B2), while the other half receive treatment B2 followed by treatment B1 (sequence B2B1). This is a two-group crossover design.

Crossover designs assume that a long enough period elapses between measurements so that the effects of one treatment are *washed out* before the next treatment is applied. This is known as the assumption of no *carryover* effects.

When a crossover design is analyzed using repeated measures, the interaction is the only term of interest. The *F*-test of the between factor tests whether means across each sequence are equal—a test of secondary interest. The *F*-test of the within factor tests whether the response is different across the time periods—also of secondary interest. The *F*-test for interaction tests whether the change in response across time is the same for both sequences. The interaction can only be significant if the treatments affect the outcome differently. Hence, to specify a crossover design requires the careful specification of the interaction effects.

With this background, we present an example. Suppose researchers want to investigate the reduction in heart-beat rate caused by the administration of a certain drug using a simple two-period crossover design. The researchers want a sample size large enough to detect a drop in heart-beat rate from 95 to 90 with a power of 90% at the 0.05 significance level. Previous studies have shown a within-patient autocorrelation of 0.50 and a standard deviation of 3.98.

The hypothesized interaction is calculated using the Standard Deviation of Means Calculator tool. This window is loaded using the *Sm* icon to the right of the 'Number of Factors' box, or by selecting 'Means to Sm Estimator' from the Tools menu.

Once this tool is displayed, enter the four treatment heart-beat rates forming a 2 by 2 matrix as follows:

95 90

90 95

The required standard deviation of the interaction effects will be displayed at the bottom of the window as the value of Sm(AB), which is 2.5.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Sample Size**
Based on.. **F-Test**
Term **B1*W1**
F-Test .... **Checked**
Minimum Power **0.90**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type.. **Means, Interaction Effects**
Factor B1 Levels.. **2**
Factor B1 Means .... **List of Means**
Factor B1 List of Means.. **90 95**
Factor W1 Levels.. **2**
Factor W1 Means .... **List of Means**
Factor W1 List of Means.. **90 95**
K's (Multipliers) **1.0**
Group Allocation .... **Equal (n1 = n2 = ···)**

Interactions Tab

B1*W1 Interaction Enter Effect As .... **Std Dev of Effects**
B1*W1 Interaction Std Dev of Effects **2.5**

Variances Tab

Input Type.. **Constant σ and ρ**
σ (Standard Deviation).. **3.98**
ρ (Autocorrelation) **0.5**

Output

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

Design Report

Solve For: Sample Size Based on the Regular F-Test (Term: B1*W1)

			Sample Size						
Term*	Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (2)	F	0.5224	5	10	1	2.5	3.45	0.725	0.05
W1 (2)	F	0.9338	5	10	1	2.5	1.99	1.256	0.05
B1*W1	F	0.9338	5	10	1	2.5	1.99	1.256	0.05

* The numbers in parentheses represent the number of levels associated with the factor.

We only display the interaction term since that is the only term of interest. We note that 90% power is achieved when *n* is 5. This corresponds to a total sample size of 10 subjects.

Example 5 – Power of a Completed Cross-Over Design

The following analysis of variance table was generated by **NCSS** for a set of crossover data. Find the power of the interaction *F*-test assuming a significance level of 0.05.

Analysis of Variance Table

Model Term	DF	Sum of Squares	Mean Square	F-Ratio	P-Value
A: Sequence	1	89397.6	89397.6	1.19	0.285442
B(A): Subject	28	2110739	75383.54		
C: Period	1	117395.3	117395.3	1.40	0.246854
AC	1	122401.7	122401.7	1.46	0.237263
BC(A)	28	2349752	83919.72		
S	0				
Total (Adjusted)	59	4789686			
Total	60				

Means and Standard Errors

Term	Count	Mean	Standard Error
All	60	492.2000	
A: Sequence			
1	30	453.6000	50.12768
2	30	530.8000	50.12768
C: Period			
1	30	447.9667	52.88973
2	30	536.4333	52.88973
AC: Sequence, Period			
1, 1	15	364.2000	74.79738
1, 2	15	543.0000	74.79738
2, 1	15	531.7333	74.79738
2, 2	15	529.8666	74.79738

One difficulty in analyzing an existing crossover design is determining an appropriate value for the hypothesized interaction effects. One method is to find the standard deviation of the interaction effects by taking the square root of the Sum of Squares for the interaction divided by the total number of observations. In this case,

$$\sigma_{Interaction} = \sqrt{\frac{122401.7}{60}}$$

$$= 45.1667$$

Another method is to find the individual interaction effects by subtraction. This method proceeds as follows:

First, subtract the Period means from the Sequence by Period means.

$$\begin{bmatrix} 364.2000 & 531.7333 \\ 543.0000 & 529.8666 \end{bmatrix} - \begin{bmatrix} 447.9667 \\ 536.4333 \end{bmatrix} = \begin{bmatrix} -83.7667 & 83.7667 \\ 6.5667 & -6.5667 \end{bmatrix}$$

Next, compute the column means and subtract them from the current values. This results in the effects.

$$\begin{bmatrix} -83.7667 & 83.7667 \\ 6.5667 & -6.5667 \end{bmatrix} - \begin{bmatrix} -38.60000 & 38.6000 \\ -38.60000 & 38.6000 \end{bmatrix} = \begin{bmatrix} -45.1667 & 45.1667 \\ 45.1667 & -45.1667 \end{bmatrix}$$

Finally, compute the standard deviation of the effects. Since the mean of the effects is zero, the standard deviation is

$$\sigma_{Interaction} = \sqrt{\frac{(-45.1667)^2 + (45.1667)^2 + (45.1667)^2 + (-45.1667)^2}{4}}$$

$$= 45.1667$$

Another difficulty that must be solved is to estimate the autocorrelation and within-subject standard deviation. From the above printout, we note that MSB = 75383.54 and MSW = 83919.72. Plugging these values into the estimating equations

$$\hat{\rho} = \frac{MSB - MSW}{MSB + (T-1)MSW}$$

$$\hat{\sigma}^2 = \frac{MSW}{1-\hat{\rho}}$$

yields

$$\hat{\rho} = \frac{75383.54 - 83919.72}{75383.54 + (2-1)83919.72} = -0.05358447$$

$$\hat{\sigma}^2 = \frac{83919.72}{1 + 0.05358447} = 79651.63$$

so that

$$\hat{\sigma} = \sqrt{79651.63} = 282.2262$$

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 5** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Power**
Geisser-Greenhouse..................................... **Checked**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type .... **Means, Interaction Effects**
Factor B1 Levels **2**
Factor B1 Means .... **List of Means**
Factor B1 List of Means **453.6 530.8**
Factor W1 Levels **2**
Factor W1 Means .... **List of Means**
Factor W1 List of Means **447.9667 536.4333**
K's (Multipliers) **1.0**
Group Allocation .... **Equal (n1 = n2 = ⋯ = n)**
n (Size Per Group)... **15**

Interactions Tab

B1*W1 Interaction Enter Effect As .... **Std Dev of Effects**
B1*W1 Interaction Std Dev of Effects **45.1667**

Variances Tab

Input Type .... **Constant σ and ρ**
σ (Standard Deviation).................................. **282.2262**
ρ (Autocorrelation) .... **-0.05358447**

Output

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

Design Report

Solve For: Power

			Sample Size						
Term*	Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (2)	GG F	0.1832	15	30	1	38.60	194.14	0.199	0.05
W1 (2)	GG F	0.2078	15	30	1	44.23	204.84	0.216	0.05
B1*W1	GG F	0.2147	15	30	1	45.17	204.84	0.220	0.05

* The numbers in parentheses represent the number of levels associated with the factor.

Notice that these power values are low. Fifteen was not a large enough sample size to detect the interaction value near 45.

Example 6 – Validation using O'Brien and Muller (1993)

O'Brien and Muller's article in the book edited by Edwards (1993) analyze the power of a two-group repeated-measures experiment in which three measurements are made on each subject.

The hypothesized means are

	Group 1	Group 2
Time 1	3	1
Time 2	12	5
Time 3	8	7

The covariance matrix is

	Time 1	Time 2	Time 3
Time 1	25	16	12
Time 2	16	64	30
Time 3	12	30	36

With *n*'s of 12, 18, and 24 and an alpha of 0.05, they obtained power values using the Wilks' Lambda test. Their reported power values are

Power Values for each Term

n	Group	Time	Interaction
12	0.326	0.983	0.461
18	0.467	0.999	0.671
24	0.589	0.999	0.814

O'Brien, in a private communication, re-ran these data using the Geisser-Greenhouse correction. His results were as follows:

Power Values for each Term

n	Group	Time	Interaction
12	0.326	0.993	0.486
18	0.467	0.999	0.685
24	0.589	0.999	0.819

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 6** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Power**
Geisser-Greenhouse.. **Checked**
Wilks' Lambda .... **Checked**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type.. **Means in Spreadsheet**
Factor B1 Levels.. **2**
Factor W1 Levels... **3**
Columns Containing the Means........................ **C1-C2**
K's (Multipliers) **1.0**
Group Allocation .... **Equal (n1 = n2 = ··· = n)**
n (Size Per Group)... **12 18 24**

Variances Tab

Input Type.. **Variance-Covariance Matrix in Spreadsheet**
Columns Containing the Covariance Matrix........ **C4-C6**

Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6
1	3	1		25	16	12
2	12	5		16	64	30
3	8	7		12	30	36

Output

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

Results for Factor B1 (Levels = 2)

Solve For: Power

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.3263	12	24	1	1.67	5.17	0.322	0.05
Wilks	0.3263	12	24	1	1.67	5.17	0.322	0.05
GG F	0.4673	18	36	1	1.67	5.17	0.322	0.05
Wilks	0.4673	18	36	1	1.67	5.17	0.322	0.05
GG F	0.5889	24	48	1	1.67	5.17	0.322	0.05
Wilks	0.5889	24	48	1	1.67	5.17	0.322	0.05

Results for Factor W1 (Levels = 3)

Solve For: Power

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.9909	12	24	1	2.86	2.73	1.047	0.05
Wilks	0.9825	12	24	1	2.86	2.73	1.047	0.05
GG F	0.9997	18	36	1	2.86	2.73	1.047	0.05
Wilks	0.9995	18	36	1	2.86	2.73	1.047	0.05
GG F	1.0000	24	48	1	2.86	2.73	1.047	0.05
Wilks	1.0000	24	48	1	2.86	2.73	1.047	0.05

Results for Term B1*W1

Solve For: Power

		Sample Size						
Test	Power	Average Group n	Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
GG F	0.4822	12	24	1	1.31	2.73	0.481	0.05
Wilks	0.4605	12	24	1	1.31	2.73	0.481	0.05
GG F	0.6810	18	36	1	1.31	2.73	0.481	0.05
Wilks	0.6706	18	36	1	1.31	2.73	0.481	0.05
GG F	0.8157	24	48	1	1.31	2.73	0.481	0.05
Wilks	0.8136	24	48	1	1.31	2.73	0.481	0.05

PASS agrees exactly with O'Brien's calculations for Wilks tests results. The results are slightly different for the Geisser-Greenhouse F-tests because the formulas used to make these calculation have been upgraded.

Example 7 – Unequal Group Sizes

Usually, in the planning stages, the group sample sizes are equal. Occasionally, however, you may want to plan for a situation in which one group will have a larger sample size than the others. Also, when doing a power analysis on a study that has already been conducted, the group sample sizes are often unequal.

In this example, we will re-analyze the Example 3. However, we will now assume that there were four subjects in group 1 and eight subjects in group 2. The setup and output for this example are as follows.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 7** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For .... **Power**
F-Test .... **Checked**
Alpha for All Terms **0.05**
Number of Factors .... **1 Between, 1 Within**
Mean Input Type .... **Means in Spreadsheet**
Factor B1 Levels **2**
Factor W1 Levels **3**
Columns Containing the Means **C1-C2**
K's (Multipliers) **1.0**
Group Allocation .... **Unequal (Enter n1, n2, ... Individually)**
n1, n2, ... (List) **4 8**

Variances Tab

Input Type .... **Constant σ and ρ**
σ (Standard Deviation) **0.70710681**
ρ (Autocorrelation) **0.16667**

Input Spreadsheet Data

Row	C1	C2
1	14.5	19
2	16.0	18
3	17.5	19

Output

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

Design Report

Solve For: Power

Term*	Test	Power	Sample Size Average Group n	Sample Size Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (2)	F	1.0000	6	12	1	1.26	0.47	2.667	0.05
W1 (3)	F	0.9986	6	12	1	0.62	0.37	1.660	0.05
B1*W1	F	0.9986	6	12	1	0.62	0.37	1.660	0.05

* The numbers in parentheses represent the number of levels associated with the factor.
n's: 4 8

Note that the values of *n* are shown to one decimal place. That is because the value reported is the average value of *n*. The actual *n*'s are shown following the report.

Example 8 – Designs with More Than Two Factors

Occasionally, you will have a design that has more than two factors. We will now show you how to compute the necessary sample size for such a design.

Suppose your design calls for two between-subject factors, Age (A) and Gender (G), and two within-subject factors, dose-level (D) and application method (M). The number of levels of these four factors are, respectively, 3, 2, 4, and 2.

Our first task is to determine appropriate effect values for each of the terms. We decide to ignore the interactions during the planning and only consider the factors themselves. The desired difference to be detected among the three age groups can be represented by the means 80, 88, and 96. The desired difference to be detected among the two genders can be represented by the means 80 and 96. The desired difference to be detected among the four dose levels is represented by the means 80, 82, 84, and 86. The desired difference to be detected among the two application methods is represented by the means 80 and 86.

Our next task is to specify the covariance matrix. From previous experience, we have found that a constant value of 20.0 is appropriate for the standard deviation.

The autocorrelation pattern is more complex because there are two types of within factors. We need to understand how the measurements will be obtained to come up with a realistic autocorrelation pattern. Suppose that the four doses are administered at four equal-spaced time points using application method 1, then again using application method 2. This suggests that an autoregressive autocorrelation pattern with $\rho = 0.7$ can be used for doses and a constant autocorrelation matrix with $\rho = 0.5$ can be used for application method. These two matrices can then be combined using a Kronecker product.

Finally, suppose we decide to calculate the power using the Geisser-Greenhouse test at the following sample sizes: 2, 4, 6, 8, 10, 20, 30, and 40.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 8** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design: Means Tab

Solve For ....**Power**
Geisser-Greenhouse....................................**Checked**
Alpha for All Terms **0.05**
Number of Factors ....**2 Between, 2 Within**
Mean Input Type...**Means, Interaction Effects**
Factor B1 Levels...**3**
Factor B1 Means ....**List of Means**
Factor B1 List of Means................................**80 88 96**
Factor B2 Levels...**2**
Factor B2 Means ....**List of Means**
Factor B2 List of Means................................**80 96**
Factor W1 Levels..**4**
Factor W1 Means ....**List of Means**
Factor W1 List of Means...............................**80 82 84 86**
Factor W2 Levels..**2**
Factor W2 Means ....**List of Means**
Factor W2 List of Means...............................**80 86**
K's (Multipliers) **1.0**
Group Allocation ....**Equal (n1 = n2 = ··· = n)**
n (Size Per Group)..**2 4 6 8 10 20**

Interactions Tab

(Use these settings for all interactions)
Interaction Enter Effect As ....**Multiple of Another Term**
Interaction Multiplier......................................**1.0**
Interaction Basis Term **W1**

Variances Tab

Input Type...**Non-Constant σ's and ρ's**
Pattern ....**Constant σ**
σ **20**
Factor W1 Autocorrelation Structure..............**AR(1)**
Factor W1 ρ **0.7**
Factor W2 Autocorrelation Structure..............**Constant**
Factor W2 ρ **0.5**

Output

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

Numeric Results

Design Report

Solve For: Power

Term*	Test	Power	Sample Size: Average Group n	Sample Size: Total N	Effect Multiplier K	Standard Deviation of Effects σm	Standard Deviation σ	Effect Size σm / σ	Alpha
B1 (3)	GG F	0.1834	2	12	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	0.3732	2	12	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.0848	2	12	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.1876	2	12	1	3.00	8.23	0.364	0.05
B1 (3)	GG F	0.4389	4	24	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	0.7387	4	24	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.2361	4	24	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.3937	4	24	1	3.00	8.23	0.364	0.05
B1 (3)	GG F	0.6438	6	36	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	0.9026	6	36	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.3979	6	36	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.5620	6	36	1	3.00	8.23	0.364	0.05
B1 (3)	GG F	0.7881	8	48	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	0.9668	8	48	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.5545	8	48	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.6937	8	48	1	3.00	8.23	0.364	0.05
B1 (3)	GG F	0.8804	10	60	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	0.9895	10	60	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.6889	10	60	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.7916	10	60	1	3.00	8.23	0.364	0.05
B1 (3)	GG F	0.9959	20	120	1	6.53	14.26	0.458	0.05
B2 (2)	GG F	1.0000	20	120	1	8.00	14.26	0.561	0.05
W1 (4)	GG F	0.9732	20	120	1	2.24	5.68	0.394	0.05
W2 (2)	GG F	0.9771	20	120	1	3.00	8.23	0.364	0.05

* The numbers in parentheses represent the number of levels associated with the factor.

Standard Deviations and Correlation Matrix

Name	W1W1	W1W2	W2W1	W2W2	W3W1	W3W2	W4W1	W4W2
W1W1	20.0000	0.5000	0.700	0.350	0.490	0.245	0.3430	0.1715
W1W2	0.5000	20.0000	0.350	0.700	0.245	0.490	0.1715	0.3430
W2W1	0.7000	0.3500	20.000	0.500	0.700	0.350	0.4900	0.2450
W2W2	0.3500	0.7000	0.500	20.000	0.350	0.700	0.2450	0.4900
W3W1	0.4900	0.2450	0.700	0.350	20.000	0.500	0.7000	0.3500
W3W2	0.2450	0.4900	0.350	0.700	0.500	20.000	0.3500	0.7000
W4W1	0.3430	0.1715	0.490	0.245	0.700	0.350	20.0000	0.5000
W4W2	0.1715	0.3430	0.245	0.490	0.350	0.700	0.5000	20.0000

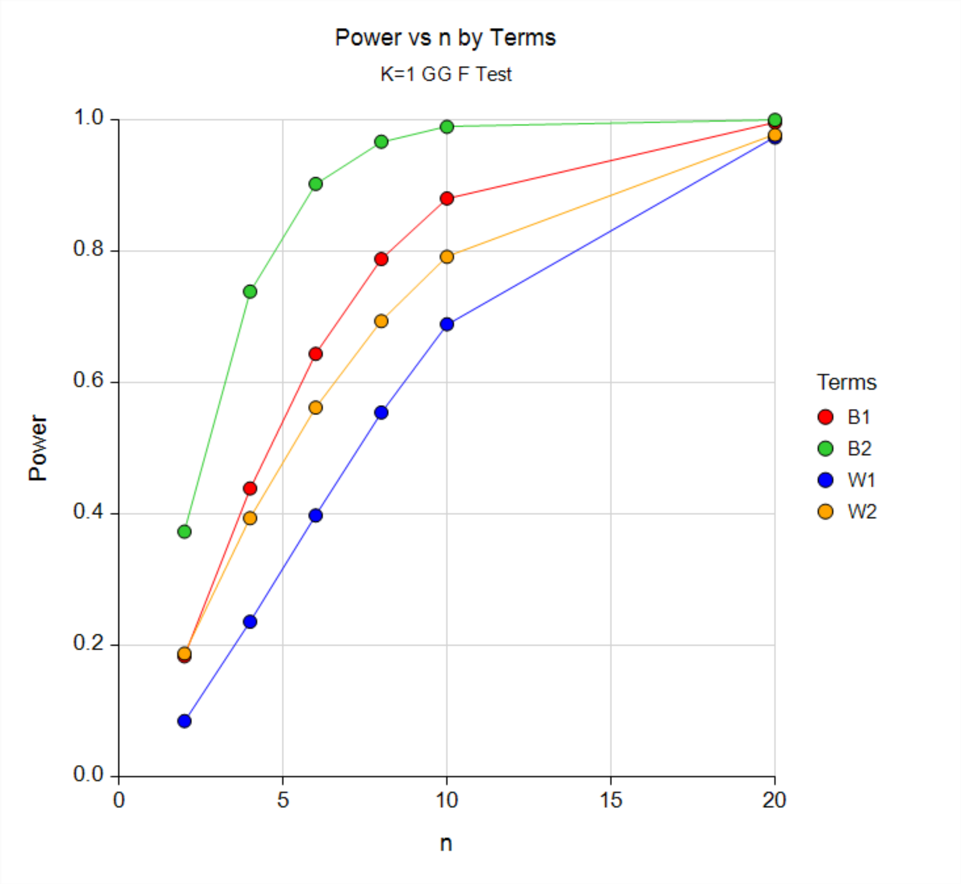
SD's on the diagonal. Correlations on the off diagonal(s).

This report gives the power values for the various terms and sample sizes that were entered. It also shows the autocorrelation pattern that resulted from the Kronecker product of the two within factor correlation matrices.

It is much easier to consider the following plot to interpret the results.

Plots Section

Plots



From this chart, we can see that the first within-subject factor, dose level, has a power much lower than the other factors.