Chapter 431

Tests for Two Means in a Repeated Measures Design

Introduction

This module calculates the power for testing the time-averaged difference (TAD) between two means in a *repeated measures* design. A repeated measures design is one in which subjects are observed repeatedly over time. Measurements may be taken at pre-determined intervals (e.g. weekly or at specified time points following the administration of a particular treatment), or at random times so there are variable intervals between repeated measurements.

Time-averaged difference analysis is often used when the outcome to be measured varies with time. For example, suppose that you want to compare two treatment groups based on the means of a certain outcome such as blood pressure. It is known that a person's blood pressure depends on several instantaneous factors such as amount of sleep, excitement level, mood, exercise, etc. If only a single measurement is taken from each patient, then the comparison of mean values from the two groups may be invalid because of the large degree of variation in blood pressure levels among patients. The precision of the experiment is increased by taking multiple measurements from each individual and comparing the time-averaged difference between the two groups. Care must be taken in the analysis because of the correlation that is introduced when several measurements are taken from the same individual. The covariance structure may take on several forms depending on the nature of the experiment and the subjects involved. This procedure allows you to calculate sample sizes using four different covariance patterns: Compound Symmetry, AR(1), Banded(1), and Simple.

This procedure can be used to calculate sample size and power for tests of pairwise contrasts in a mixed models analysis of repeated measures data. Mixed models analysis of repeated measures data is also employed to provide more flexibility in covariance specification and a greater degree of robustness in the presence of missing data, provided that the data can be assumed to be missing at random.

Technical Details

Theory and Notation

For a study with n_1 subjects in group 1 and n_2 subjects in group 2 (for a total of N subjects), each measured m times, the time-averaged difference (d) of a continuous response between two groups can be estimated using the following model:

$$y_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_{ij}, \quad i = 1, \cdots, N; \ j = 1, \cdots, m \ ,$$

where

y_{ij} is the jth response from subject i,

β_0 is the model intercept,

β_1 is the treatment effect or the time-averaged difference between groups 1 and 2 (i.e., $\beta_1 = d$),

x_i is a binary group assignment variable, which is equal to 1 if the ith subject is in group 1 and equal to 0 if the ith subject is in group 2, and

ε_{ij} is the normal, random error associated with the observation y_{ij}.

Accounting for the relationship between repeated measurements, the model presented above can be written in matrix form as

$$\boldsymbol{y}_i = \boldsymbol{X}_i'\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i,$$

where

$\boldsymbol{y}_i = (y_{i1} \quad y_{i2} \quad \cdots \quad y_{im})'$ is an $m \times 1$ vector of responses from subject i,

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}_{m \times 2} \quad \text{if the } i^{\text{th}} \text{ subject is in group 1,}$$

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{pmatrix}_{m \times 2} \quad \text{if the } i^{\text{th}} \text{ subject is in group 2,}$$

$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$ is the vector of model parameters, and

$\boldsymbol{\varepsilon}_i \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{R}_i)$ is the vector of correlated random errors for the observations from subject i,

and

> $\text{var}(y_{ij}) = \sigma^2$ is the residual variance for a single observation, and $\mathbf{R}_i$ is the $m \times m$ common correlation matrix for all subjects. The contents of $\mathbf{R}_i$ depend on the assumed within-subject correlation structure.

We can stack the data in a single vector and matrix form as follows:

$$\mathbf{y} = (\mathbf{y}_1', \mathbf{y}_2', \ldots, \mathbf{y}_N')'$$

$$\mathbf{X} = (\mathbf{X}_1, '\mathbf{X}_2,' \cdots, X_N')'$$

$$\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', \ldots, \boldsymbol{\varepsilon}_N')'$$

and the model for the N equations can be compressed into one as

$$\boldsymbol{y} = \boldsymbol{X}'\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

with

$$\mathbf{V} = \text{var}(\mathbf{y})$$

$$= \sigma^2 \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_N \end{pmatrix}$$

$$= \sigma^2 \mathbf{R}$$

as the covariance (or variance - covariance) matrix.

Covariance Pattern

In a repeated measures design with N subjects, each measured m times, observations from a single subject may be correlated, and a pattern for their covariance must be specified. In this case, **V** will have a block-diagonal form:

$$V = \begin{pmatrix} V_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & V_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & V_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & V_N \end{pmatrix},$$

where $\mathbf{V}_i$ are $m \times m$ covariance matrices corresponding to the ith subject. The **0**'s represent $m \times m$ matrices of zeros giving zero covariances for observations on different subjects. This routine allows the specification of four different covariance matrix types: Compound Symmetry, AR(1), Banded(1), and Simple.

Compound Symmetry

A compound symmetry covariance model assumes that all covariances are equal, and all variances on the diagonal are equal. That is

$$\boldsymbol{V}_i = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \rho & \cdots & \rho \\ \rho & \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \rho & \cdots & 1 \end{pmatrix}_{m \times m}$$

where σ^2 is the residual variance and ρ is the correlation between observations on the same subject.

AR(1)

An AR(1) (autoregressive order 1) covariance model assumes that all variances on the diagonal are equal and that covariances t time periods apart are equal to $\sigma^2 \rho^t$. That is

$$\boldsymbol{V}_i = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{m-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{m-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{m-3} \\ \rho^3 & \rho^2 & \rho & 1 & \cdots & \rho^{m-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \rho^{m-4} & \cdots & 1 \end{pmatrix}_{m \times m}$$

where σ^2 is the residual variance and ρ is the correlation between observations on the same subject.

Banded(1)

A Banded(1) (banded order 1) covariance model assumes that all variances on the diagonal are equal, covariances for observations one time period apart are equal to $\sigma^2 \rho$, and covariances for measurements greater than one time period apart are equal to zero. That is

$$\boldsymbol{V}_i = \sigma^2 \begin{pmatrix} 1 & \rho & 0 & 0 & \cdots & 0 \\ \rho & 1 & \rho & 0 & \cdots & 0 \\ 0 & \rho & 1 & \rho & \cdots & 0 \\ 0 & 0 & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}_{m \times m}$$

where σ^2 is the residual variance and ρ is the correlation between observations on the same subject.

Simple

A simple covariance model assumes that all variances on the diagonal are equal and that all covariances are equal to zero. That is

$$
\boldsymbol{V}_i = \sigma^2 \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}_{m \times m}
$$

where σ^2 is the residual variance.

Model Estimation

With $\widehat{V} = \hat{\sigma}^2 \widehat{R}$, then estimates of the regression coefficients from the above regression model are given as

$$
\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix}
$$

$$
= \left(\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\widehat{\mathbf{V}}^{-1}\mathbf{y}
$$

and the variance of $\widehat{\boldsymbol{\beta}}$ is

$$
\mathrm{var}(\widehat{\boldsymbol{\beta}}) = \begin{pmatrix} \mathrm{var}(\hat{\beta}_0) & \mathrm{cov}(\hat{\beta}_0,\hat{\beta}_1) \\ \mathrm{cov}(\hat{\beta}_0,\hat{\beta}_1) & \mathrm{var}(\hat{\beta}_1) \end{pmatrix}
$$

$$
= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}
$$

which is estimated by substituting $\widehat{V}$ for $\mathbf{V}$.

Hypothesis Test

A two-sided test that the time-averaged difference between the two groups is equal to zero is equivalent to the test of $H_0: \beta_1 = 0$ vs. $H_1: \beta_1 \neq 0$. Similarly, the upper and lower one-sided tests are $H_0: \beta_1 \leq 0$ vs. $H_1: \beta_1 > 0$ and $H_0: \beta_1 \geq 0$ vs. $H_1: \beta_1 < 0$, respectively. The test can be carried out using the test statistic

$$z = \frac{\hat{\beta}_1}{\sqrt{\text{var}(\hat{\beta}_1)}} \to N(0,1).$$

If the standard deviation is unknown and estimated, a *t* test should be used. In practice, this test is often carried out by calculating the average response for each individual and then using a two-sample *t* test. If the data are balanced, the test can also be carried out in **NCSS** using Repeated Measures GLM and specifying a comparison such as "Each with First". In the case where the data are not balanced, the test could be carried out using SAS® `PROC MIXED` or SAS® `PROC GLM`. In both cases a `REPEATED` statement should be used, along with a statement such as

```
ESTIMATE 'A-B' treat 1 -1;   or   LSMEANS treat/ PDIFF;
```

Power Calculations

Sample sizes for repeated measures studies are often calculated as if a simple trial with no repeated measures was planned, which results in a higher calculated sample size than would be found if the correlation between repeated measures were taken into consideration. With an idea of the correct covariance structure, and an estimate of the within-patient correlation, you can get a better estimate of the power and sample size necessary to achieve your objectives. If you have no indication of the correct covariance structure for the experiment, then the compound symmetry (program default) is likely to be adequate. If you have no previous estimate of the within-patient correlation, then Brown and Prescott (2006) suggest using a conservative prediction of the correlation, i.e., a higher correlation than anticipated.

For a two-sided test where it is assumed that $d > 0$ (without loss of generality),

$$\text{Power} = 1 - \beta = \Pr(\text{rejecting } H_0 | H_1)$$

$$= \Pr\left(\left| \frac{\hat{\beta}_1}{\sqrt{\text{var}(\hat{\beta}_1)}} \right| > z_{1-\alpha/2} | H_1 \right)$$

$$\approx \Pr\left(\frac{\hat{\beta}_1}{\sqrt{\text{var}(\hat{\beta}_1)}} > z_{1-\alpha/2} | H_1 \right) \text{ since it is assumed that } d > 0$$

$$= \Pr\left(\frac{\hat{\beta}_1 - d}{\sqrt{\text{var}(\hat{\beta}_1)}} > z_{1-\alpha/2} - \frac{d}{\sqrt{\text{var}(\hat{\beta}_1)}} | H_1 \right)$$

$$= 1 - \Phi\left(z_{1-\alpha/2} - \frac{d}{\sqrt{\mathrm{var}(\hat{\beta}_1)}}\right),$$

where $\Phi()$ is the standard normal density function, and α and β are the probabilities of type I and type II error, respectively. For a one-sided test, α is used in place of $\alpha/2$.

Since a *t* test is usually used to test for a group difference in a case such as this, we should note here that the power calculation using the standard normal distribution represents an approximation to the actual power achieved by the *t* test. We feel that it might be more appropriate to use the non-central *t* distribution; however, since the calculation is based on numerous assumptions about the covariance structure that influence the results, it seems unnecessary to worry about the small gain in precision that may occur by using the non-central *t* distribution. For this reason, along with the fact that this is the published method, we have elected to follow the methods of Brown and Prescott (2006), Liu and Wu (2005), Diggle et al. (1994) and use the standard normal distribution in power and sample size calculations.

Calculating Power for Testing Pairwise Contrasts of Fixed Effects in Mixed Models

Mixed Model Theory and Notation

A linear mixed model incorporates both fixed and random effects. Fixed effects are those effects in the model whose values are assumed constant, or unchanging. Random effects are those effects in the model that are assumed to have arisen from a distribution, resulting in another source of random variation other than residual variation. Brown and Prescott (2006) demonstrates how this methodology may be used to calculate the sample size and power for testing pairwise contrasts of fixed effects in a mixed models analysis of repeated measures data. For an experiment with N subjects, p fixed effect parameters, and q random effect parameters, the general mixed model can be expressed using matrix notation as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \cdots, N$$

where

$\boldsymbol{y}_i$ is an $n_i \times 1$ vector of responses for subject *i*,

$\boldsymbol{X}_i$ is an $n_i \times p$, full-rank design matrix of fixed effects for subject *i*,

$\boldsymbol{\beta}$ is an $p \times 1$ vector of fixed effects parameters,

$\boldsymbol{Z}_i$ is an $n_i \times q$ design matrix of the random effects for subject *i*,

$\boldsymbol{u}_i$ is a $q \times 1$ vector of random effects for subject *i* which has means of zero and scaled covariance matrix **G**,

$\boldsymbol{\varepsilon}_i$ is an $n_i \times 1$ vector of errors for subject *i* with zero mean and scaled covariance $\boldsymbol{\Sigma}_i$.

The covariance of $\boldsymbol{y}_i$, $\mathrm{var}(\boldsymbol{y}_i) = \boldsymbol{V}_i$, can be written as

$$V_i = \mathrm{var}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i)$$

$$= \mathbf{Z}\mathrm{var}(\mathbf{u}_i)\mathbf{Z}' + \mathrm{var}(\boldsymbol{\varepsilon}_i)$$

$$= \mathbf{Z}_i\mathbf{G}\mathbf{Z}'_i + \boldsymbol{\Sigma}_i.$$

We can stack the data in a single vector and matrix form as follows:

$$\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)'$$

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_N)'$$

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_N \end{pmatrix}$$

$$\boldsymbol{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N)'$$

$$\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N)'$$

and the mixed model for the *N* equations can be compressed into one as

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon},$$

with

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{V}_N \end{pmatrix} = \boldsymbol{ZGZ'} + \boldsymbol{\Sigma}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_N \end{pmatrix}$$

is the covariance (or variance-covariance) matrix.

Mixed Model Estimation

Estimates of the variance components **G** and $\boldsymbol{\Sigma}$ are found using maximum likelihood (ML) or restricted/residual maximum likelihood (REML) methods. From these estimates, $\widehat{\boldsymbol{G}}$ and $\widehat{\boldsymbol{\Sigma}}$, an estimate of **V** is obtained as $\widehat{\boldsymbol{V}} = \boldsymbol{Z}\widehat{\boldsymbol{G}}\boldsymbol{Z}' + \widehat{\boldsymbol{\Sigma}}$. The fixed effects are then estimated as

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\widehat{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widehat{\boldsymbol{V}}^{-1}\boldsymbol{y}$$

with the variance of $\widehat{\boldsymbol{\beta}}$ estimated as

$$\text{var}(\widehat{\boldsymbol{\beta}}) = (\boldsymbol{X}'\widehat{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}.$$

If $\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{R}$, then these estimation equations are identical to the TAD estimation equations presented earlier, except for the fact that $\boldsymbol{\beta}$ may contain more than two parameters, i.e., a parameter for each fixed effect being modeled. In the TAD model presented above, β_1 represents the difference between two treatment means, d. In the mixed model formulation presented here, β_1, β_2, etc. represent individual treatment effects. If there are no random effects, then we can use this routine for TAD to calculate the approximate power for testing pairwise contrasts of fixed effects in mixed models designs.

Brown and Prescott (2006) presents an example on page 228 of an experiment for which the power for testing pairwise contrasts can be calculated using this procedure. To determine the relative efficacy of three treatments in controlling hypertension, patients are assigned to one of the three treatments and blood pressure is measured at four follow-up visits. The study aims to determine the differences in average blood pressure among the three treatments.

Testing Fixed Effects

Significance tests for fixed effects can be done using tests based on the *t* distribution. We can define tests of fixed and random effects as contrasts

$$\boldsymbol{C} = \boldsymbol{L}'\widehat{\boldsymbol{\beta}} = \boldsymbol{0},$$

respectively. For example, in a trial containing three treatments A, B, and C, a pairwise comparison of treatments A and C is given by the contrast

$$\boldsymbol{C}_{AC} = \boldsymbol{L}'\widehat{\boldsymbol{\beta}} = (0 \quad 1 \quad 0 \quad -1)\widehat{\boldsymbol{\beta}} = \hat{\beta}_A - \hat{\beta}_C,$$

where the first term in $\boldsymbol{\beta}$ is the intercept term, and the other three terms are the treatment effects.

For a single comparison, the test statistic is given by

$$t_{df} = \frac{\boldsymbol{L}'\widehat{\boldsymbol{\beta}}}{\text{SE}(\boldsymbol{L}'\widehat{\boldsymbol{\beta}})}$$

$$= \frac{\hat{\beta}_j - \hat{\beta}_h}{SE(\hat{\beta}_j - \hat{\beta}_h)},$$

where *df* is the degrees of freedom, usually determined using the Satterthwaite approximation, and $\hat{\beta}_j$ and $\hat{\beta}_h$ $(j \neq h)$ are estimated treatment effects.

Contrasts such as this can be tested in SAS® using the ESTIMATE statement or by including the PDIFF option in an LSMEANS statement. For example, if the variable designating three treatments, A, B, and C, were called "treat", then I could use the following statements in PROC MIXED to test for a difference between A and C

```
ESTIMATE 'A-C' treat 1 0 –1;
```

or

```
LSMEANS treat/ PDIFF;
```

The latter statement would produce tests of all pairwise comparisons of the levels of the treatment variable. The former would only test the difference between groups A and C. Of course, these comparison statements must be used in conjunction with appropriate model and class statements (see pages 233-237 of Brown and Prescott (2006) for an example analyzed using SAS® PROC MIXED).

Estimates of the correlation (ρ) and the standard deviation (σ) for use in power calculations can be found using SAS® PROC MIXED. For a model fit using compound symmetry, σ^2 and ρ can be estimated as the sum of the variance parameters, and the compound symmetry variance parameter divided by the sum of the variance parameters, respectively. For AR(1), Banded(1), and Simple covariance models, σ^2 and ρ can be estimated as the residual variance, and the correlation between adjacent measurements, respectively. Alternatively, the R and RCORR options may be used within the REPEATED statement to display the covariance and correlation matrices, from which the parameter estimates can be determined.

Example 1 – Determining Power

Researchers are planning a study of the impact of a new drug on heart rate. They want to evaluate the time-averaged difference in heart rate between subjects who take the new drug, and subjects who take the standard drug. 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 between the two treatments.

Similar studies have found an average heart rate of 93 for individuals taking the standard drug, a standard deviation of 9, 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 83.7, for a difference of 9.3. The test will be conducted at the 0.05 significance level.

What power does the study achieve over a range of possible sample sizes?

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 Tab
_____

Solve For ....**Power**
Alternative Hypothesis ....**Two-Sided**
Alpha..**0.05**
Group Allocation ....**Equal (N1 = N2)**
Sample Size Per Group ....**4 to 20 by 2**
D1 ....**9.3**
M..**4**
Covariance Type...**AR(1)**
Sigma...**9**
Rho **0.7**

Output

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

Numeric Reports

Numeric Results
_____

Solve For: Power
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: AR(1)
_____

Power	Sample Size			Number of of Time Points M	Difference D1	Standard Deviation of a Single Observation Sigma	Autocorrelation Rho	Alpha
	N1	N2	N					
0.42660	4	4	8	4	9.3	9	0.7	0.05
0.58468	6	6	12	4	9.3	9	0.7	0.05
0.70890	8	8	16	4	9.3	9	0.7	0.05
0.80135	10	10	20	4	9.3	9	0.7	0.05
0.86742	12	12	24	4	9.3	9	0.7	0.05
0.91318	14	14	28	4	9.3	9	0.7	0.05
0.94407	16	16	32	4	9.3	9	0.7	0.05
0.96448	18	18	36	4	9.3	9	0.7	0.05
0.97773	20	20	40	4	9.3	9	0.7	0.05

Power   The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1 and N2   The number of items sampled from each population.
N   The total sample size. N = N1 + N2.
M   The number of time points (repeated measurements) at which each subject is observed.
D1   The difference between the means of groups 1 and 2 under the alternative hypothesis.
Sigma   The standard deviation of a single observation. It is the same for both groups.
Rho   The correlation between any two observations on the same subject.
Alpha   The probability of rejecting a true null hypothesis.

Summary Statements
_____

A parallel two-group repeated-measures design with 4 measurements on each subject will be used to test whether
the Group 1 mean is different from the Group 2 mean. The comparison will be made using a two-sided, two-sample
time-averaged difference in means test, with a Type I error rate (α) of 0.05. The covariance structure for repeated
observations on the same subject is AR(1), with a correlation between observations on the same subject
(autocorrelation) of 0.7. To detect a difference of 9.3, with sample sizes of 4 for Group 1 and 4 for Group 2, the
power is 0.4266.
_____

Dropout-Inflated Sample Size

Dropout Rate	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
	N1	N2	N	N1'	N2'	N'	D1	D2	D
20%	4	4	8	5	5	10	1	1	2
20%	6	6	12	8	8	16	2	2	4
20%	8	8	16	10	10	20	2	2	4
20%	10	10	20	13	13	26	3	3	6
20%	12	12	24	15	15	30	3	3	6
20%	14	14	28	18	18	36	4	4	8
20%	16	16	32	20	20	40	4	4	8
20%	18	18	36	23	23	46	5	5	10
20%	20	20	40	25	25	50	5	5	10

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.

N1, N2, and N   The evaluable sample sizes at which power is computed (as entered by the user). If N1 and N2 subjects are evaluated out of the N1' and N2' subjects that are enrolled in the study, the design will achieve the stated power.

N1', N2', and N'   The number of subjects that should be enrolled in the study in order to obtain N1, N2, and N evaluable subjects, based on the assumed dropout rate. N1' and N2' are calculated by inflating N1 and N2 using the formulas N1' = N1 / (1 - DR) and N2' = N2 / (1 - DR), with N1' and N2' 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.)

D1, D2, and D   The expected number of dropouts. D1 = N1' - N1, D2 = N2' - N2, and D = D1 + D2.

Dropout Summary Statements

Anticipating a 20% dropout rate, 5 subjects should be enrolled in Group 1, and 5 in Group 2, to obtain final group sample sizes of 4 and 4, respectively.

References

Brown, H. and Prescott, R., 2006. Applied Mixed Models in Medicine. 2nd ed. John Wiley & Sons Ltd. Chichester, West Sussex, England. Chapter 6.

Liu, H. and Wu, T., 2005. 'Sample Size Calculation and Power Analysis of Time-Averaged Difference.' Journal of Modern Applied Statistical Methods, Vol. 4, No. 2, pages 434-445.

Diggle, P.J., Liang, K.Y., and Zeger, S.L., 1994. Analysis of Longitudinal Data. Oxford University Press. New York, New York. Chapter 2.

This report gives the power for each value of the other parameters.

Plots Section

Plots
_____
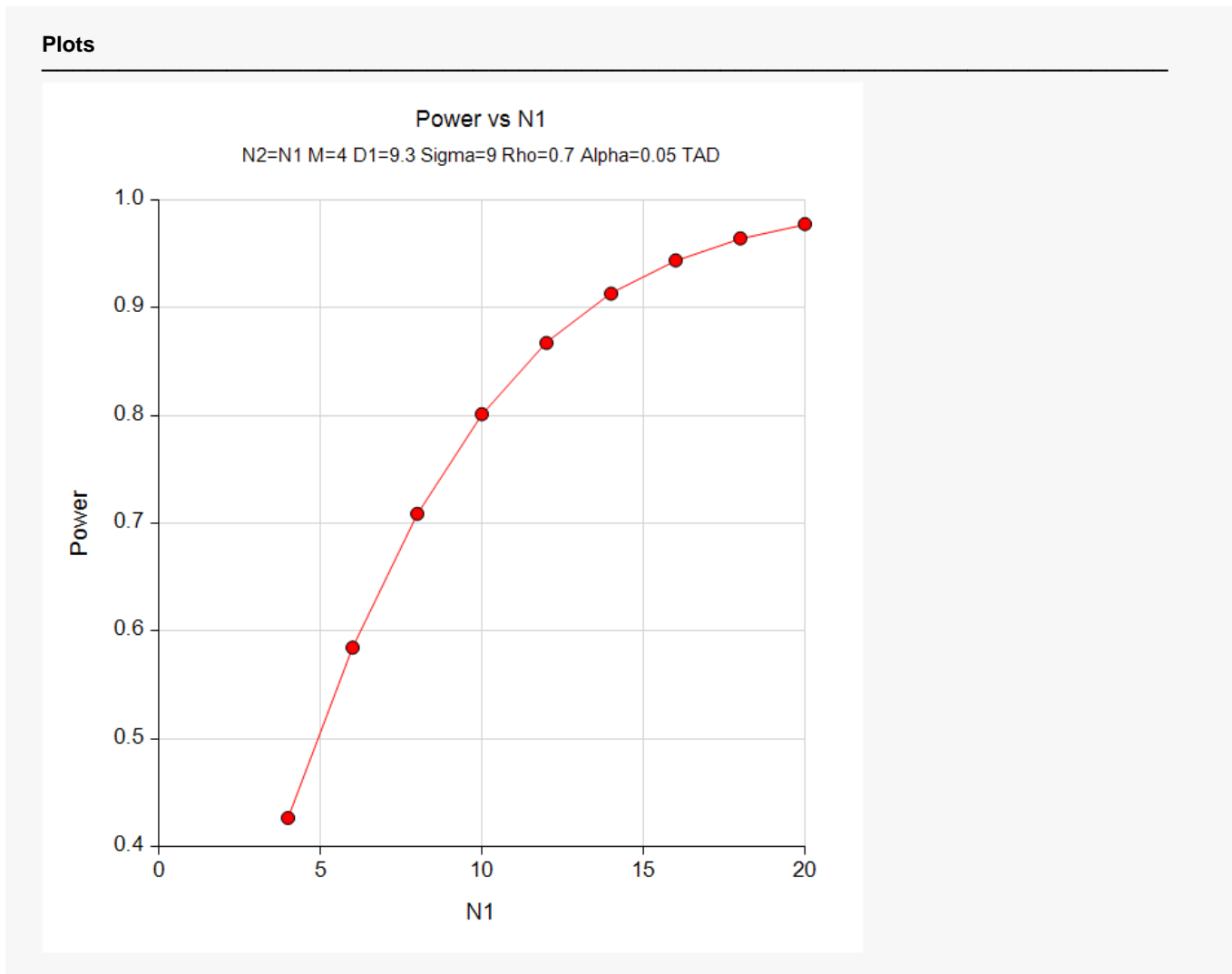
The chart shows the relationship between power and N1 when the other parameters in the design are held constant.

Example 2 – Finding the Sample Size

Continuing with Example 1, the researchers want to determine the exact sample size necessary to achieve at least 80% power.

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 Tab

Solve For ....**Sample Size**
Alternative Hypothesis ....**Two-Sided**
Power...**0.80**
Alpha..**0.05**
Group Allocation ....**Equal (N1 = N2)**
D1 ....**9.3**
M..**4**
Covariance Type...**AR(1)**
Sigma...**9**
Rho **0.7**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: AR(1)

Power		Sample Size			Number of of Time Points	Difference	Standard Deviation of a Single Observation	Autocorrelation	
Target	Actual	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.8	0.80135	10	10	20	4	9.3	9	0.7	0.05

A group sample size of 10 is required to achieve at least 80% power.

Example 3 – Varying the Difference Between the Means

Continuing with Examples 1 and 2, the researchers want to evaluate the impact on power of varying the size of the difference between the means for a range of sample sizes from 2 to 8 per group.

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 Tab

Solve For ....**Power**
Alternative Hypothesis ....**Two-Sided**
Alpha...**0.05**
Group Allocation ....**Equal (N1 = N2)**
Sample Size Per Group ....**4 to 8 by 1**
D1 ....**4 to 11 by 1**
M..**4**
Covariance Type...**AR(1)**
Sigma..**9**
Rho **0.7**

Output

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

Numeric Results

Solve For: Power
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: AR(1)

| | Sample Size | | | Number of of Time Points | Difference | Standard Deviation of a Single Observation | Autocorrelation | |
Power	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.11574	4	4	8	4	4	9	0.7	0.05
0.15728	4	4	8	4	5	9	0.7	0.05
0.20758	4	4	8	4	6	9	0.7	0.05
0.26631	4	4	8	4	7	9	0.7	0.05
0.33245	4	4	8	4	8	9	0.7	0.05
0.40428	4	4	8	4	9	9	0.7	0.05
0.47949	4	4	8	4	10	9	0.7	0.05
0.55544	4	4	8	4	11	9	0.7	0.05
0.13427	5	5	10	4	4	9	0.7	0.05
0.18591	5	5	10	4	5	9	0.7	0.05
0.24835	5	5	10	4	6	9	0.7	0.05
0.32050	5	5	10	4	7	9	0.7	0.05
0.40016	5	5	10	4	8	9	0.7	0.05
0.48423	5	5	10	4	9	9	0.7	0.05
0.56900	5	5	10	4	10	9	0.7	0.05
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

Plots

These plots show how the power depends on the difference to be detected, D1, as well as the group sample size, N1.

Example 4 – Impact of the Number of Repeated Measurements

Continuing with Examples 1 - 3, the researchers want to study the impact on the sample size if they change the number of measurements made on each individual. Their experimental protocol calls for four measurements that are 30 minutes apart. They want to see the impact of taking twice that many measurements.

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 Tab

Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided**
Power...	**0.80**
Alpha...	**0.05**
Group Allocation	**Equal (N1 = N2)**
D1	**9.3**
M..	**4 8**
Covariance Type...	**AR(1)**
Sigma...	**9**
Rho	**0.7**

Output

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

Numeric Results
_____
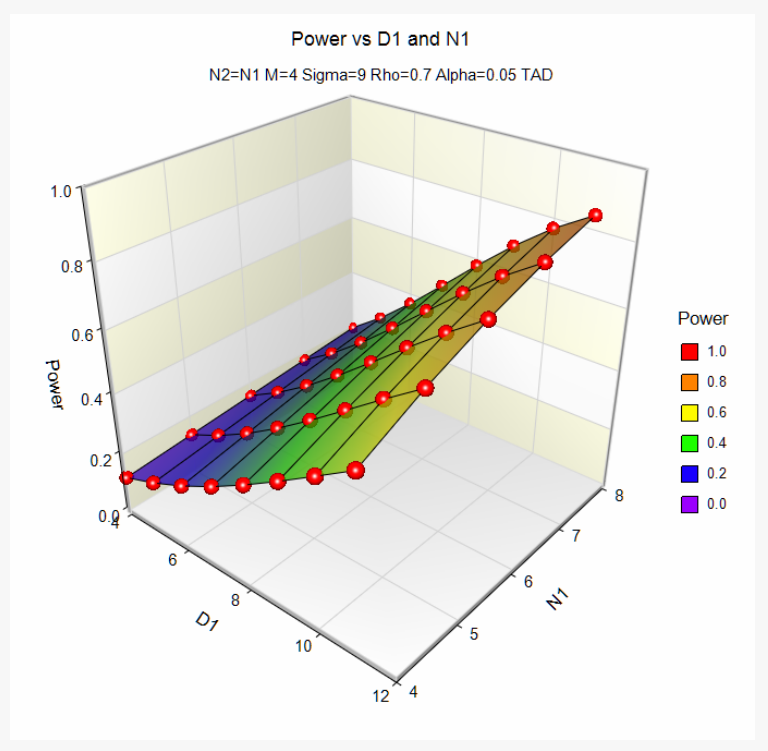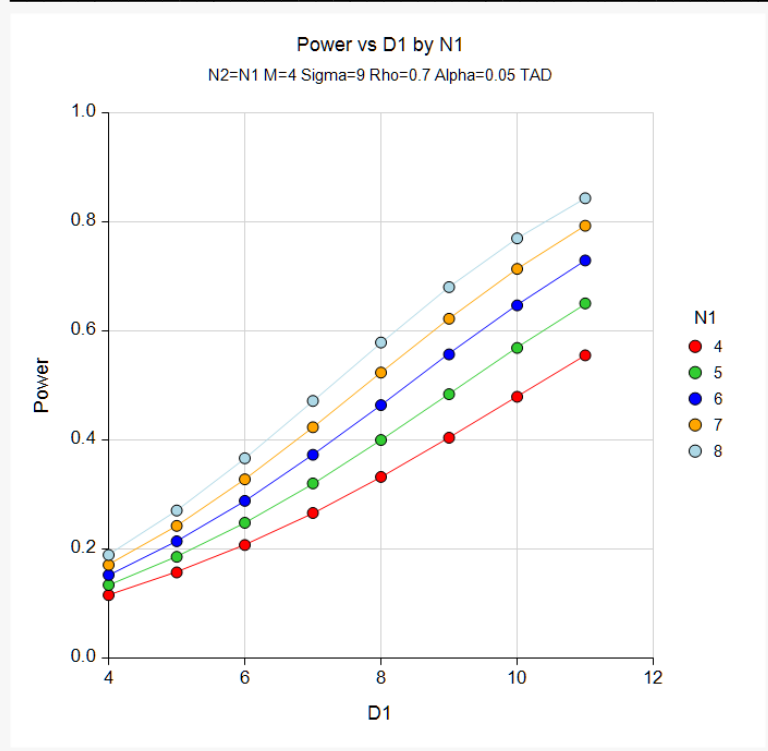Solve For: Sample Size
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: AR(1)
_____

Power		Sample Size			Number of of Time Points	Difference	Standard Deviation of a Single Observation	Autocorrelation	
Target	Actual	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.8	0.80135	10	10	20	4	9.3	9	0.7	0.05
0.8	0.84737	8	8	16	8	9.3	9	0.7	0.05

Doubling the number of repeated measurements per individual decreases the group sample size by 2. This reduction in sample size may not justify the additional four measurements on each subject.

Example 5 – Validation using Diggle et al. (1994)

Diggle et al. (1994) page 31 presents an example of calculating the sample size for a TAD study. They calculate the group sample sizes for the cases where d/σ ranges from 0.2 to 0.5, ρ ranges from 0.2 to 0.8, alpha = 0.05, M = 3, and power = 0.8. Note that Diggle et al. (1994) uses a one-sided test.

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 Tab

Solve For ....**Sample Size**
Alternative Hypothesis ....**One-Sided**
Power...**0.80**
Alpha..**0.05**
Group Allocation ....**Equal (N1 = N2)**
D1 ....**0.2 to 0.5 by 0.1**
M..**3**
Covariance Type...**Compound Symmetry**
Sigma...**1**
Rho **0.2 0.5 0.8**

Output

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

Numeric Results
_____

Solve For: Sample Size
Hypotheses: One-Sided. H0: D = 0   vs.   H1: D > 0 (or D < 0)
Covariance Type: Compound Symmetry
_____

| Power | | Sample Size | | | Number of of Time Points | Difference | Standard Deviation of a Single Observation | Autocorrelation | |
Target	Actual	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.8	0.80178	145	145	290	3	0.2	1	0.2	0.05
0.8	0.80154	207	207	414	3	0.2	1	0.5	0.05
0.8	0.80012	268	268	536	3	0.2	1	0.8	0.05
0.8	0.80475	65	65	130	3	0.3	1	0.2	0.05
0.8	0.80154	92	92	184	3	0.3	1	0.5	0.05
0.8	0.80270	120	120	240	3	0.3	1	0.8	0.05
0.8	0.80885	37	37	74	3	0.4	1	0.2	0.05
0.8	0.80321	52	52	104	3	0.4	1	0.5	0.05
0.8	0.80012	67	67	134	3	0.4	1	0.8	0.05
0.8	0.81343	24	24	48	3	0.5	1	0.2	0.05
0.8	0.80028	33	33	66	3	0.5	1	0.5	0.05
0.8	0.80109	43	43	86	3	0.5	1	0.8	0.05

_____

The sample sizes calculated by **PASS** match the results of Diggle et al. (1994) very closely, with slight differences due to rounding. If you calculate the sample sizes by hand, using the formula given in Diggle et al. (1994), page 31, your answers will match those of **PASS**.

Example 6 – Validation of Sample Size Calculation for Mixed Models Analysis using Brown and Prescott (2006)

Brown and Prescott (2006) pages 268 and 269 presents an example of calculating the sample size for pairwise contrasts in a hypertension trial to by analyzed using mixed models. The analysis of repeated DBP measurements from four post-treatment visits using a compound symmetry covariance pattern resulted in the following covariance matrix for each subject:

$$V_i = 76 \begin{pmatrix} 1 & 0.53 & 0.53 & 0.53 \\ 0.53 & 1 & 0.53 & 0.53 \\ 0.53 & 0.53 & 1 & 0.53 \\ 0.53 & 0.53 & 0.53 & 1 \end{pmatrix}$$

From this matrix they determine that $\rho = 0.53$ and $\sigma^2 = 76$ ($\sigma = 8.718$).

The trial followed several hundred patients given one of three treatments. Brown and Prescott calculate the group sample size to be 31 for a future study involving four post-treatment visits to detect a difference in DBP of 5 mmHg at the 5% significance level with 80% power.

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 Tab

Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided**
Power..	**0.80**
Alpha...	**0.05**
Group Allocation	**Equal (N1 = N2)**
D1	**5**
M..	**4**
Covariance Type...	**Compound Symmetry**
Sigma..	**8.718**
Rho	**0.53**

Output

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

Numeric Report for M = 4

Numeric Results
_____

Solve For: Sample Size
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: Compound Symmetry
_____

Power		Sample Size			Number of of Time Points	Difference	Standard Deviation of a Single Observation	Autocorrelation	
Target	Actual	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.8	0.80125	31	31	62	4	5	8.718	0.53	0.05

The sample size of 31 calculated by **PASS** matches the results of Brown and Prescott (2006) exactly.

Brown and Prescott further calculate the sample size for the case where no account is taken of repeated measurements and the case of 10 repeated measurements. If we change the number of repeated measurements to 1 and 10, we get the following output (**Example 6b** settings file):

Numeric Report for M = 1, 10

Numeric Results
_____

Solve For: Sample Size
Hypotheses: Two-Sided. H0: D = 0   vs.   H1: D ≠ 0
Covariance Type: Compound Symmetry
_____

Power		Sample Size			Number of of Time Points	Difference	Standard Deviation of a Single Observation	Autocorrelation	
Target	Actual	N1	N2	N	M	D1	Sigma	Rho	Alpha
0.8	0.80226	48	48	96	1	5	8.718	0.53	0.05
0.8	0.80651	28	28	56	10	5	8.718	0.53	0.05

In both cases, the results of **PASS** match those of Brown and Prescott (2006) exactly.