Chapter 470

Equivalence Tests for the Ratio of Two Means

Introduction

This procedure calculates power and sample size of statistical tests for *equivalence* tests from parallel-group design with two groups. This routine deals with the case in which the statistical hypotheses are expressed in terms of mean ratios rather than mean differences.

The details of testing the equivalence of two treatments using a parallel-group design are given in the chapter entitled "Equivalence Tests for Two Means using Differences" and will not be repeated here. If the logarithms of the responses can be assumed to follow a normal distribution, hypotheses about equivalence in terms of the ratio can be transformed into hypotheses about the difference. The details of this analysis are given in Julious (2004).

Equivalence Testing Using Ratios

It will be convenient to adopt the following specialize notation for the discussion of these tests.

Parameter	PASS Input/Output	Interpretation
μ_T	Not used	*Treatment mean*. This is the treatment mean.
μ_R	Not used	*Reference mean*. This is the mean of a reference population.
ϕ_L, ϕ_U	RL, RU	*Margin of equivalence*. These limits define an interval of the ratio of the means in which their difference is so small that it may be ignored.
ϕ	R1	*True ratio*. This is the value of $\phi = \mu_T / \mu_R$ at which the power is calculated.

Note that the actual values of μ_T and μ_R are not needed. Only their ratio is needed for power and sample size calculations.

The null hypothesis of non-equivalence is

$$H_0: \phi \le \phi_L \text{ or } \phi \ge \phi_U \text{ where } \phi_L < 1, \phi_U > 1.$$

and the alternative hypothesis of equivalence is

$$H_1: \phi_L < \phi < \phi_U$$

Log-Transformation

In many cases, hypotheses stated in terms of ratios are more convenient than hypotheses stated in terms of differences. This is because ratios can be interpreted as scale-less percentages, but differences must be interpreted as actual amounts in their original scale. Hence, it has become a common practice to take the following steps in hypothesis testing.

1. State the statistical hypotheses in terms of ratios.

2. Transform these into hypotheses about differences by taking logarithms.

3. Analyze the logged data—that is, do the analysis in terms of the difference.

4. Draw the conclusion in terms of the ratio.

The details of step 2 for the null hypothesis are as follows.

$$\phi_L \leq \phi \leq \phi_U$$

$$\Rightarrow \phi_L \leq \left\{ \frac{\mu_T}{\mu_R} \right\} \leq \phi_U$$

$$\Rightarrow \ln(\phi_L) \leq \left\{ \ln(\mu_T) - \ln(\mu_R) \right\} \leq \ln(\phi_U)$$

Thus, a hypothesis about the ratio of the means on the original scale can be translated into a hypothesis about the difference of two means on the logged scale.

When performing an equivalence test on the difference between means, the usual procedure is to set the equivalence limits symmetrically above and below zero. Thus the equivalence limits will be plus or minus an appropriate amount. The common practice is to do the same when the data are being analyzed on the log scale. However, when symmetric limits are set on the log scale, they do not translate to symmetric limits on the original scale. Instead, they translate to limits that are the inverses of each other.

Perhaps these concepts can best be understood by considering an example. Suppose the researchers have determined that the lower equivalence limit should be 80% on the original scale. Since they are planning to use a log scale for their analysis, they transform this limit to the log scale by taking the logarithm of 0.80. The result is -0.223144. Wanting symmetric limits, they set the upper equivalence limit to 0.223144. Exponentiating this value, they find that exp(0.223144) = 1.25. Note that 1/(0.80) = 1.25. Thus, the limits on the original scale are 80% and 125%, not 80% and 120%.

Using this procedure, appropriate equivalence limits for the ratio of two means can be easily determined. Here are a few sets of equivalence limits.

Specified Percent Change	Lower Limit Original Scale	Upper Limit Original Scale	Lower Limit Log Scale	Upper Limit Log Scale
-25%	75.0%	133.3%	-0.287682	0.287682
+25%	80.0%	125.0%	-0.223144	0.223144
-20%	80.0%	125.0%	-0.223144	0.223144
+20%	83.3%	120.0%	-0.182322	0.182322
-10%	90.0%	111.1%	-0.105361	0.105361
+10%	90.9%	110.0%	-0.095310	0.095310

Note that negative percent-change values specify the lower limit first, while positive percent-change values specify the upper limit first. After the first limit is found, the other limit is calculated as its inverse.

Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean. This parameter can be used to represent the variation in the data because of a unique relationship that it has in the case of log-normal data.

Suppose the variable X is the logarithm of the original variable Y. That is, $X = \ln(Y)$ and $Y = \exp(X)$. Label the mean and variance of X as μ_X and σ_X^2, respectively. Similarly, label the mean and variance of Y as μ_Y and σ_Y^2, respectively. If X is normally distributed, then Y is log-normally distributed. Julious (2004) presents the following well-known relationships between these two variables

$$\mu_Y = \left(e^{\mu_X + \frac{\sigma_X^2}{2}} \right)$$

$$\sigma_Y^2 = \mu_Y^2 \left(e^{\sigma_X^2} - 1 \right)$$

From this relationship, the coefficient of variation of Y can be expressed as

$$COV_Y = \frac{\sqrt{\mu_Y^2 \left(e^{\sigma_X^2} - 1 \right)}}{\mu_Y}$$

$$= \sqrt{e^{\sigma_X^2} - 1}$$

Solving this relationship for σ_X^2, the standard deviation of X can be stated in terms of the coefficient of variation of Y. This equation is

$$\sigma_X = \sqrt{\ln\left(COV_Y^2 + 1 \right)}$$

Similarly, the mean of X is

$$\mu_X = \frac{\mu_Y}{\ln\left(COV_Y^2 + 1 \right)}$$

One final note: for parallel-group designs, σ_X^2 equals σ_d^2, the average variance used in the t-test of the logged data.

Thus, the hypotheses can be stated in the original (Y) scale and then the power can be analyzed in the transformed (X) scale.

Power Calculation

As is shown above, the hypotheses can be stated in the original (Y) scale using ratios or the logged (X) scale using differences. Either way, the power and sample size calculations are made using the formulas for testing the equivalence of the difference in two means. These formulas are presented another chapter and are not duplicated here.

Procedure Options

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

Design Tab

The Design tab contains the parameters associated with this test such as the means, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. Under most situations, you will select either *Power* or *Sample Size (N1)*.

Power and Alpha

Power

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

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

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

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when you reject the null hypothesis of non-equivalent means when in fact the means are nonequivalent.

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

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

You can enter a range of values such as *0.05, 0.10, 0.15* or *0.05 to 0.15 by 0.01*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on *N1* or *N2* or both.

The options are

- **Equal (N1 = N2)**

    This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N2, solve for N1**

    Select this option when you wish to fix *N2* at some value (or values), and then solve only for *N1*. Please note that for some values of *N2*, there may not be a value of *N1* that is large enough to obtain the desired power.

- **Enter R = N2/N1, solve for N1 and N2**

    For this choice, you set a value for the ratio of *N2* to *N1*, and then PASS determines the needed *N1* and *N2*, with this ratio, to obtain the desired power. An equivalent representation of the ratio, *R*, is

    $$N2 = R * N1.$$

- **Enter percentage in Group 1, solve for N1 and N2**

    For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed *N1* and *N2* with this percentage to obtain the desired power.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of *N2* to *N1*. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of *N2* to *N1* while solving for *N1* and *N2*. Only sample size combinations with this ratio are considered.

N2 is related to *N1* by the formula:

$$N2 = [R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1 (e.g., *N1* = 10 and *N2* = 20, or *N1* = 50 and *N2* = 100).

R must be greater than 0. If *R* < 1, then *N2* will be less than *N1*; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for *N1* and *N2*. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When <u>Not</u> Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**

  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**

  This choice permits you to enter different values for *N1* and *N2*.

- **Enter N1 and R, where N2 = R * N1**

  Choose this option to specify a value (or values) for *N1*, and obtain *N2* as a ratio (multiple) of *N1*.

- **Enter total sample size and percentage in Group 1**

  Choose this option to specify a value (or values) for the total sample size (*N*), obtain *N1* as a percentage of *N*, and then *N2* as *N - N1*.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal (N1 = N2)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for *N1*, and also the value for *N2*.

The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where N2 = R * N1."

N1 is the number of items or individuals sampled from the Group 1 population.

N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

N2 is the number of items or individuals sampled from the Group 2 population.

N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where N2 = R * N1."*

R is the ratio of *N2* to *N1*. That is,

$$R = N2/N1$$

Use this value to obtain *N2* as a multiple (or proportion) of *N1*.

N2 is calculated from *N1* using the formula:

$$N2=[R \times N1],$$

where the value *[Y]* is the next integer $\geq Y$.

For example, setting *R* = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If *R* < 1, then *N2* will be less than N1; if *R* > 1, then *N2* will be greater than *N1*. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines *N1* and *N2*.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size – Equivalence Limits

RU (Upper Equivalence Limit)

Enter the upper equivalence limit for the ratio of the two means. When the ratio of the means is between this value and RL, the two means are said to be equivalent. The value must be greater than one. A popular choice is 1.25. Note that this value is not a percentage.

If you enter *1/RL*, then 1/RL will be calculated and used here. This choice is commonly used because RL and 1/RL give limits that are of equal magnitude on the log scale.

RL (Lower Equivalence Limit)

Enter the lower equivalence limit for the ratio of the two means. When the ratio of the means is between this value and RU, the two means are said to be equivalent. The value must be less than one. A popular choice is 0.80. Note that this value is not a percentage.

If you enter *1/RU*, then 1/RU will be calculated and used here. This choice is commonly used because RU and 1/RU give limits that are of equal magnitude on the log scale.

Effect Size – True Ratio

R1 (True Ratio)

This is the value of the ratio of the two means at which the power is to be calculated. Usually, the ratio will be assumed to be one. However, some authors recommend calculating the power using a ratio of 1.05 since this will require a larger sample size.

Effect Size – Coefficient of Variation

COV (Coefficient of Variation)

The coefficient of variation is used to specify the variability (standard deviation). It is important to realize that this is the COV defined on the original (not logged) scale. This value must be determined from past experience or from a pilot study. It is most easily calculated from the within mean-square error of the analysis of variance of the logged data using the relationship

$$COV_Y = \sqrt{e^{\sigma_w^2} - 1} \,.$$

If prior studies used a t-test to analyze the logged data, you will not have a direct estimate of $\hat{\sigma}_w^2$. However, the two variances, σ_d^2 and σ_w^2, are functionally related. The relationship between these quantities is $\sigma_d^2 = 2\sigma_w^2$.

Example 1 – Finding Power

A company has developed a generic drug for treating rheumatism and wants to show that it is equivalent to the standard drug. A parallel-group design will be used to test the equivalence of the two drugs.

Researchers have decided to set the lower limit of equivalence at 0.80. Past experience leads the researchers to set the COV to 1.50. The significance level is 0.05. The power will be computed assuming that the true ratio is either 1.00 or 1.05. Sample sizes between 50 and 550 will be included in the analysis.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Equivalence Tests for the Ratio of Two Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Ratio of Two Means**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Power**
Alpha..	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**50 to 550 by 100**
RU (Upper Equivalence Limit)	**1/RL**
RL (Lower Equivalence Limit).................	**0.80**
R1 (True Ratio)	**1.0 1.05**
COV (Coefficient of Variation)	**1.50**

Annotated Output

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

Numeric Results for Testing Equivalence Using a Parallel-Group Design

Power	N1	N2	N	R1	COV	Lower Equiv. Limit	Upper Equiv. Limit	Alpha
0.00000	50	50	100	1.00	1.50	0.80	1.25	0.050
0.10488	150	150	300	1.00	1.50	0.80	1.25	0.050
0.48431	250	250	500	1.00	1.50	0.80	1.25	0.050
0.71704	350	350	700	1.00	1.50	0.80	1.25	0.050
0.84945	450	450	900	1.00	1.50	0.80	1.25	0.050
0.92208	550	550	1100	1.00	1.50	0.80	1.25	0.050
0.00000	50	50	100	1.05	1.50	0.80	1.25	0.050
0.09731	150	150	300	1.05	1.50	0.80	1.25	0.050
0.43421	250	250	500	1.05	1.50	0.80	1.25	0.050
0.63661	350	350	700	1.05	1.50	0.80	1.25	0.050
0.76018	450	450	900	1.05	1.50	0.80	1.25	0.050
0.83961	550	550	1100	1.05	1.50	0.80	1.25	0.050

Equivalence Tests for the Ratio of Two Means

Report Definitions
Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
R1 is the true ratio of the means.
COV is the coefficient of variation on the original scale.
The lower and upper equivalence limits are the maximum allowable ratios that still result in equivalence.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements
An equivalence test of means using two one-sided tests on data from a parallel-group design
with sample sizes of 50 in the reference group and 50 in the treatment group achieves 0% power
at a 5% significance level when the true ratio of the means is 1.00, the coefficient of
variation on the original, unlogged scale is 1.50, and the equivalence limits of the mean ratio
are 0.80 and 1.25.

This report shows the power for the indicated scenarios.

Plot Section
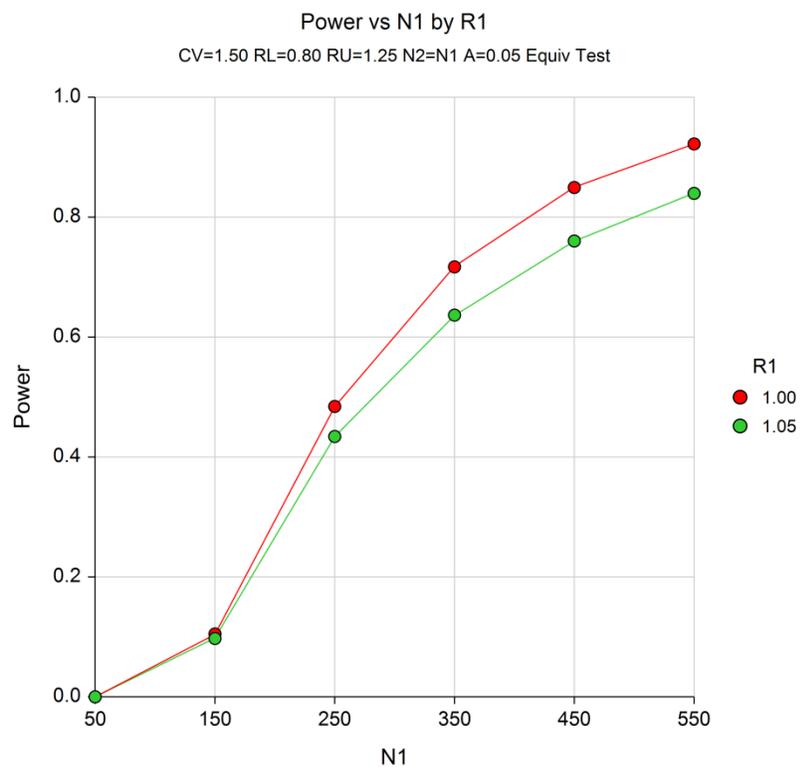
Equivalence Tests for the Ratio of Two Means
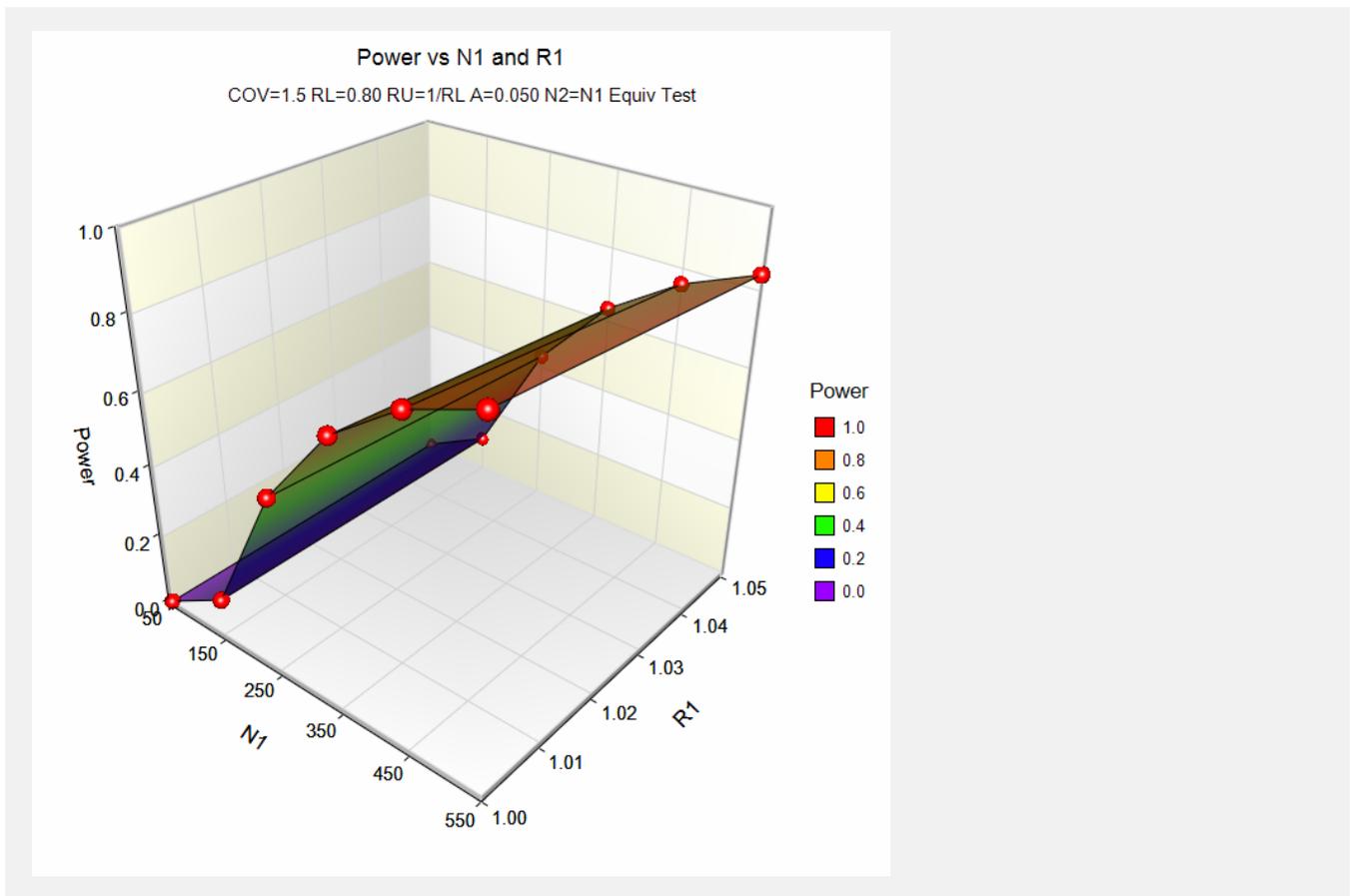
These plots shows the power versus the sample size for two values of R1.

Example 2 – Validation using Julious (2004)

Julious (2004) page 1971 presents an example of determining the sample size for a parallel-group design in which the actual ratio is 1.0, the coefficient of variation is 0.80, the equivalence limits are 0.80 and 1.25, the power is 90%, and the significance level is 0.05. He calculates the per group sample size to be 216.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Equivalence Tests for the Ratio of Two Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Ratio of Two Means**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Power..	**0.90**
Alpha...	**0.05**
Group Allocation	**Equal (N1 = N2)**
RU (Upper Equivalence Limit)	**1/RL**
RL (Lower Equivalence Limit).................	**0.80**
R1 (True Ratio)	**1.0**
COV (Coefficient of Variation)	**0.80**

Output

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

Numeric Results

Numeric Results for Testing Equivalence Using a Parallel-Group Design

Target Power	Actual Power	N1	N2	N	R1	COV	Lower Equiv. Limit	Upper Equiv. Limit	Alpha
0.90	0.90044	216	216	432	1.00	0.80	0.80	1.25	0.050

PASS also calculates the per group sample size to be 216, which matches Julious' result.