

Chapter 525

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

Introduction

This procedure calculates power and sample size of statistical tests of equivalence of the means from a 2x2 cross-over design which is analyzed with a t-test. This routine deals with the case in which the statistical hypotheses are expressed in terms mean of ratios rather than mean differences.

The details of testing the equivalence of two treatments using data from a 2x2 cross-over design are given in another chapter and will not be repeated here. If the logarithms of the responses can be assumed to follow the normal distribution, hypotheses about the equivalence of two means stated in terms of the ratio can be transformed into hypotheses about the difference. The details of this analysis are given in Julious (2004). They will only be summarized here.

Equivalence Testing Using Ratios

PASS follows the *two one-sided tests* approach described by Schuirmann (1987) and Phillips (1990). It will be convenient to adopt the following specialized notation for the discussion of these tests.

<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
μ_T	Not used	<i>Treatment mean.</i> This is the treatment mean.
μ_R	Not used	<i>Reference mean.</i> This is the mean of a reference population.
ϕ_L, ϕ_U	RL, RU	<i>Margin of equivalence.</i> These limits that define an interval of the ratio of the means in which their difference is so small that it may be ignored.
ϕ	R1	<i>True ratio.</i> 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.

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

The null hypothesis of non-equivalence is

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

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.

$$\begin{aligned} \phi_L &\leq \phi \leq \phi_U \\ \Rightarrow \phi_L &\leq \left\{ \frac{\mu_T}{\mu_R} \right\} \leq \phi_U \\ \Rightarrow \ln(\phi_L) &\leq \{ \ln(\mu_T) - \ln(\mu_R) \} \leq \ln(\phi_U) \end{aligned}$$

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.

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

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

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

Similarly, the mean of X is

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

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 in 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* for a power analysis or *Sample Size* for sample size determination.

Power and Alpha

Power

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

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 unequal means when in fact the means are unequal.

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*.

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

Sample Size

N (Total Sample Size)

This option specifies one or more values of the sample size, the number of individuals in the study (total subjects in both sequences). This value must be an integer greater than one.

When N is even, it is split evenly between the two sequences. When N is odd, the first sequence has one more subject than the second sequence.

Note that you may enter a list of values using the syntax *50,100,150,200,250* or *50 to 250 by 50*.

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 using 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 opened a new manufacturing plant and wants to show that the drug produced in the new plant is equivalent to that produced in an older plant. A cross-over design will be used to test the equivalence of drugs produced at the two plants.

Researchers have decided to set the equivalence limits for the ratio at 0.90 and 1.111 (note that $1.111 = 1/0.90$). Past experience leads the researchers to set the COV to 0.50. The significance level is 0.05. The power will be computed assuming that the true ratio is one. 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 in a 2x2 Cross-Over Design** procedure window by expanding **Means**, then **Cross-Over (2x2) Design**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design**. 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alpha.....	0.05
N (Total Sample Size).....	50 to 550 by 100
RU (Upper Equivalence Limit)	1/RL
RL (Lower Equivalence Limit).....	0.90
R1 (True Ratio).....	1.0
COV (Coefficient of Variation)	0.50

Annotated Output

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

Numeric Results

Numeric Results for Testing Equivalence Using a Cross-Over Design

Power	Total Sample Size (N)	Lower Equiv. Limit of Ratio (RL)	Upper Equiv. Limit of Ratio (RU)	True Ratio (R1)	Coefficient of Variation (COV)	Alpha	Beta
0.0000	50	0.9000	1.1111	1.0000	0.5000	0.0500	1.0000
0.2190	150	0.9000	1.1111	1.0000	0.5000	0.0500	0.7810
0.6002	250	0.9000	1.1111	1.0000	0.5000	0.0500	0.3998
0.8064	350	0.9000	1.1111	1.0000	0.5000	0.0500	0.1936
0.9101	450	0.9000	1.1111	1.0000	0.5000	0.0500	0.0899
0.9596	550	0.9000	1.1111	1.0000	0.5000	0.0500	0.0404

Report Definitions

- Power is the probability of rejecting non-equivalence when the means are equivalent.
- N is the total number of subjects split between both sequences.
- RU & RL are the upper and lower equivalence limits. Ratios between these limits are equivalent.
- R1 is the ratio of the means at which the power is computed.
- COV is the coefficient of variation on the original scale.
- Alpha is the probability of rejecting non-equivalence when the means are non-equivalent.
- Beta is the probability of accepting non-equivalence when the means are equivalent.

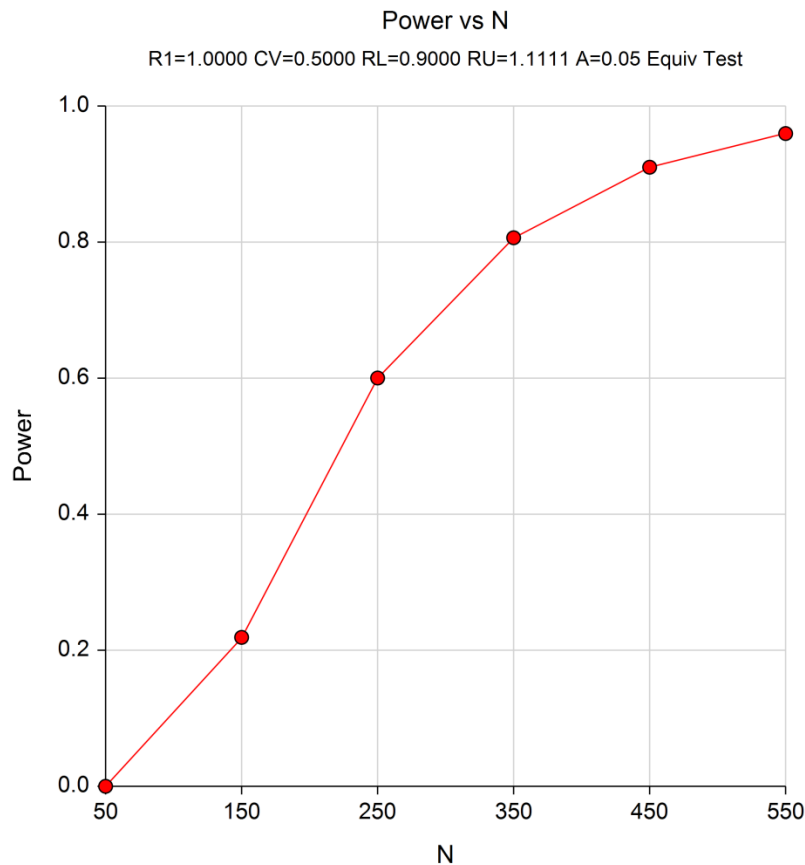
Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design

Summary Statements

In an equivalence test of means using two one-sided tests on data from a two-period cross-over design, a total sample size of 50 achieves 0% power at a 5% significance level when the true ratio of the means is 1.0000, the coefficient of variation on the original, unlogged scale is 0.5000, and the equivalence limits of the mean ratio are 0.9000 and 1.1111.

This report shows the power for the indicated scenarios. Note that if they want 90% power, they will require a sample of around 450 subjects.

Plots Section



This plot shows the power versus the sample size.

Example 2 – Validation using Julious (2004)

Julious (2004) page 1963 presents a table of sample sizes for various parameter values. The power is 0.90 and the significance level is 0.05. The COV is set to 0.25, the 'level of bioequivalence' is set to 10%, 15%, 20%, and 25%, and the true ratio is set to 1.00, the necessary sample sizes are 120, 52, 28, and 18. Note that the level of bioequivalence as defined in Julious (2004) is equal to $1 - RL$.

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 in a 2x2 Cross-Over Design** procedure window by expanding **Means**, then **Cross-Over (2x2) Design**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design**. 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Power.....	0.90
Alpha.....	0.05
RU (Upper Equivalence Limit)	1/RL
RL (Lower Equivalence Limit).....	0.90 0.85 0.80 0.75
R1 (True Ratio)	1.00
COV (Coefficient of Variation)	0.25

Output

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

Numeric Results

Numeric Results for Testing Equivalence Using a Cross-Over Design							
Power	Total Sample Size (N)	Lower Equiv. Limit of Ratio (RL)	Upper Equiv. Limit of Ratio (RU)	True Ratio (R1)	Coefficient of Variation (COV)	Alpha	Beta
0.9121	18	0.7500	1.3333	1.0000	0.2500	0.0500	0.0879
0.9023	28	0.8000	1.2500	1.0000	0.2500	0.0500	0.0977
0.9060	52	0.8500	1.1765	1.0000	0.2500	0.0500	0.0940
0.9012	120	0.9000	1.1111	1.0000	0.2500	0.0500	0.0988

Note that PASS obtains the same samples sizes as Julious (2004).