Chapter 525

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design (Log-Normal Data)

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.

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.
R_L, R_U	RL, RU	*Equivalence Limits.* 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	*Actual 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 the ratio of these values is needed for power and sample size calculations.

With $R_L < 1$ and $R_U > 1$, the null hypothesis of non-equivalence is

$$H_0: \phi \leq R_L \text{ or } \phi \geq R_U.$$

The alternative hypothesis of equivalence is

$$H_1: R_L < \phi < R_U.$$

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 alternative hypothesis are as follows:

$$H_1: R_L < \phi < R_U \quad \Rightarrow \quad H_1: R_L < \frac{\mu_T}{\mu_R} < R_U \quad \Rightarrow \quad H_1: \ln(R_L) < \ln(\mu_T) - \ln(\mu_R) < \ln(R_U)$$

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 is 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

$$\mu_Y = e^{\mu_X + \frac{\sigma_X^2}{2}}$$

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

From this relationship, the coefficient of variation of *Y* can be found to be

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

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

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

where σ_W^2 is the within mean square error from the analysis of variance of the logged data. Solving this relationship for σ_X^2, the standard deviation of *X* can be stated in terms of the coefficient of variation of *Y* as

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

Similarly, the mean of *X* is

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

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

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

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**
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 (Actual Ratio)	**1.0**
COV (Coefficient of Variation)	**0.50**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Ratio: R = Treatment Mean / Reference Mean
Hypotheses: H0: R ≤ EL or R ≥ EU vs. H1: RL < R < RU

Power	Total Sample Size N	Equivalence Limits Lower RL	Equivalence Limits Upper RU	Actual Ratio R1	COV	Alpha
0.00001	50	0.9	1.111	1	0.5	0.05
0.21897	150	0.9	1.111	1	0.5	0.05
0.60022	250	0.9	1.111	1	0.5	0.05
0.80639	350	0.9	1.111	1	0.5	0.05
0.91006	450	0.9	1.111	1	0.5	0.05
0.95957	550	0.9	1.111	1	0.5	0.05

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

Summary Statements

A 2×2 cross-over design will be used to test whether the treatment mean (μ_T) is equivalent to the reference mean (μ_R), with mean ratio equivalence limits of 0.9 and 1.111 (H0: R ≤ 0.9 or R ≥ 1.111 versus H1: 0.9 < R < 1.111, R = μ_T / μ_R). The comparison will be made using two one-sided t-tests using a log-transformation, with an overall Type I error rate (α) of 0.05. The coefficient of variation on the original scale is assumed to be 0.5. To detect a ratio of means (μ_T / μ_R) of 1, with a total sample size of 50 (allocated equally to the two sequences), the power is 0.00001.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	50	63	13
20%	150	188	38
20%	250	313	63
20%	350	438	88
20%	450	563	113
20%	550	688	138

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.
N	The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N'	The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' 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.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 63 subjects should be enrolled to obtain a final sample size of 50 subjects.

References

Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.

Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.

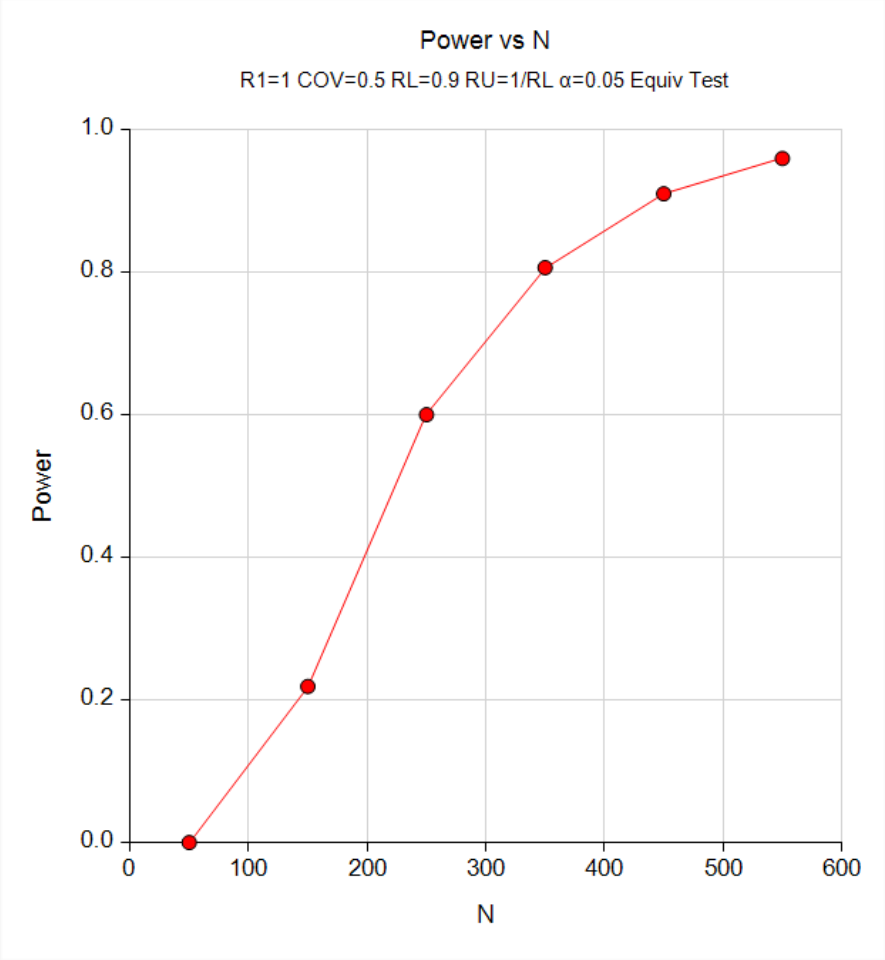
Julious, Steven A. 2004. 'Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data.' Statistics in Medicine, 23:1921-1986.

Senn, Stephen. 2002. Cross-over Trials in Clinical Research. Second Edition. John Wiley & Sons. New York.

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

Plots



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

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**
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 (Actual Ratio)	**1.0**
COV (Coefficient of Variation)	**0.25**

Output

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

Numeric Results

Solve For: Sample Size
Ratio: R = Treatment Mean / Reference Mean
Hypotheses: H0: R ≤ EL or R ≥ EU vs. H1: RL < R < RU

Power	Total Sample Size N	Equivalence Limits Lower RL	Equivalence Limits Upper RU	Actual Ratio R1	COV	Alpha
0.91211	18	0.75	1.333	1	0.25	0.05
0.90226	28	0.80	1.250	1	0.25	0.05
0.90601	52	0.85	1.176	1	0.25	0.05
0.90119	120	0.90	1.111	1	0.25	0.05

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