Chapter 129

Equivalence Tests for the Ratio of Two Means (Normal Data)

Introduction

This procedure calculates power and sample size of statistical tests for *equivalence* tests from parallel-group design with two groups when *the data are assumed to follow the normal distribution* (so the log transformation is not used). 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 this analysis are given in Hauschke *et al.* (1999) and Kieser and Hauschke (1999).

Note that when the data follow a log-normal distribution rather than the normal distribution, you should use another **PASS** procedure entitled *Equivalence Tests for the Ratio of Two Means (Log-Normal Data)* to obtain more accurate results.

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 entered directly	*Treatment mean*. This is the treatment mean.
μ_R	Not entered directly	*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

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

The alternative hypothesis of equivalence is

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

Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean of the control group. This parameter is used to represent the variation in the data. That is, $COV = \frac{\sigma}{\mu_C}$.

Power Calculation

An exact method for the calculation of power and sample size for testing equivalence based on the mean ratio when the data are assumed to be normally distributed (untransformed) was presented by Hauschke et al. (1999). This computation requires the evaluation of the noncentral bivariate t-distribution. Kieser and Hauschke (1999) present a close approximation for the case when RL = 1 / RU (the equivalence limits are the inverse of each other). This approximation is used in **PASS**. Kieser and Hauschke (1999) show that the approximation is very close to the exact method.

Model and Test Details

Suppose a comparison is to be made between two groups: a treatment (T) and a control (C). The response of interest is assumed to follow the normal distribution with (possibly different) means μ_T and μ_C and common variance σ^2. To carry out the comparison, a random sample of $N_1 = N_2$ subjects is to be obtained from each group. The parameters of the study will be presented in terms of the mean ratio μ_T/μ_C.

The equivalence hypotheses are

$$H_0: R1 \leq RL \text{ or } R1 \geq RU \quad \text{versus} \quad H_1: RL \leq R1 \leq RU$$

where $R1 = \frac{\mu_T}{\mu_C}$.

The null hypothesis H_0 is rejected in favor of the alternative if a two-sided $100(1-2\alpha)\%$ Fieller confidence interval is included completely between *RL* and *RU*.

Power Approximation

As mentioned above, **PASS** uses the approximate formulas given by Kieser and Hauschke (1999). These formulas are stated in terms of the group sample size *N1*, but they can be rearranged to give formulas for power as well.

If $R1 = 1$, use

$$N1 \geq (1 + RL^2)\left(t_{\alpha,2N1-2} + t_{\beta,2N-2}\right)^2 \left(\frac{COV}{1-RL}\right)^2$$

If $RL < R1 < 1$, use

$$N1 \geq (1 + RL^2)\left(t_{\alpha,2N1-2} + t_{\beta,2N-2}\right)^2 \left(\frac{COV}{R1-RL}\right)^2$$

If $1 < R1 < RU$, use

$$N1 \geq (1 + RU^2)\left(t_{\alpha,2N1-2} + t_{\beta,2N-2}\right)^2 \left(\frac{COV}{R1 - RU}\right)^2$$

Note that here, $t_{\alpha,2N1-2}$ denotes the $100(1-\alpha)\%$ percentile of the central t distribution with 2*N* – 2 degrees of freedom. Also, β is the probability of a type II error.

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.

Following standard procedure, researchers set the lower limit of equivalence at 0.80 and the upper limit to 1.25. 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 0.9, 0.95, 1.0, 1.05, or 1.1. Sample sizes between 500 and 2500 will be investigated.

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**
Sample Size Per Group	**500 1000 1500 2000 2500**
Equivalence Limit Input Type	**Enter Upper Equivalence Limit RU (RL = 1 / RU)**
RU (Upper Equivalence Limit)	**1.25**
R1 (Actual Ratio)	**0.9 0.95 1 1.05 1.1**
COV (Coefficient of Variation)	**1.5**

Output

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

Numeric Results

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

				Mean Ratio				
	Sample Size			Equivalence Limits				
Power	N1	N2	N	Lower RL	Upper RU	Actual R1	Coefficient of Variation COV	Alpha
0.31538	500	500	1000	0.8	1.25	0.90	1.5	0.05
0.50054	1000	1000	2000	0.8	1.25	0.90	1.5	0.05
0.64480	1500	1500	3000	0.8	1.25	0.90	1.5	0.05
0.75275	2000	2000	4000	0.8	1.25	0.90	1.5	0.05
0.83096	2500	2500	5000	0.8	1.25	0.90	1.5	0.05
0.54030	500	500	1000	0.8	1.25	0.95	1.5	0.05
0.79511	1000	1000	2000	0.8	1.25	0.95	1.5	0.05
0.91607	1500	1500	3000	0.8	1.25	0.95	1.5	0.05
0.96761	2000	2000	4000	0.8	1.25	0.95	1.5	0.05

0.98805	2500	2500	5000	0.8	1.25	0.95	1.5	0.05
0.50552	500	500	1000	0.8	1.25	1.00	1.5	0.05
0.90049	1000	1000	2000	0.8	1.25	1.00	1.5	0.05
0.98302	1500	1500	3000	0.8	1.25	1.00	1.5	0.05
0.99740	2000	2000	4000	0.8	1.25	1.00	1.5	0.05
0.99963	2500	2500	5000	0.8	1.25	1.00	1.5	0.05
0.58612	500	500	1000	0.8	1.25	1.05	1.5	0.05
0.83863	1000	1000	2000	0.8	1.25	1.05	1.5	0.05
0.94301	1500	1500	3000	0.8	1.25	1.05	1.5	0.05
0.98121	2000	2000	4000	0.8	1.25	1.05	1.5	0.05
0.99411	2500	2500	5000	0.8	1.25	1.05	1.5	0.05
0.40210	500	500	1000	0.8	1.25	1.10	1.5	0.05
0.62951	1000	1000	2000	0.8	1.25	1.10	1.5	0.05
0.78068	1500	1500	3000	0.8	1.25	1.10	1.5	0.05
0.87466	2000	2000	4000	0.8	1.25	1.10	1.5	0.05
0.93035	2500	2500	5000	0.8	1.25	1.10	1.5	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.
RL and RU	The lower and upper equivalence limits, respectively, and are the maximum allowable ratios that still result in equivalence.
R1	The actual ratio of the means at which power is calculated.
COV	The coefficient of variation on the original scale.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design will be used to test whether the two means are equivalent based on the mean ratio (R = μ1 / μ2 = μTrt / μRef), with lower and upper equivalence limits 0.8 and 1.25, respectively (H0: R ≤ 0.8 or R ≥ 1.25 versus H1: 0.8 < R < 1.25). The comparison will be made using the original (untransformed) data with two one-sided, two-sample equal-variance tests, with a Type I error rate (α) of 0.05. These tests assume that the original data in each group follow a Normal distribution with (possibly) differing means and common variance. The coefficient of variation (on the original, unlogged scale) of the reference group (σ / μ2, or σ / μRef) is assumed to be 1.5. To detect a mean ratio of 0.9 with sample sizes of 500 for Group 1 (treatment) and 500 for Group 2 (reference), the power is 0.31538.

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	**N1**	**N2**	**N**	**N1'**	**N2'**	**N'**	**D1**	**D2**	**D**
20%	500	500	1000	625	625	1250	125	125	250
20%	1000	1000	2000	1250	1250	2500	250	250	500
20%	1500	1500	3000	1875	1875	3750	375	375	750
20%	2000	2000	4000	2500	2500	5000	500	500	1000
20%	2500	2500	5000	3125	3125	6250	625	625	1250

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, 625 subjects should be enrolled in Group 1, and 625 in Group 2, to obtain final group sample sizes of 500 and 500, respectively.

References

Kieser, M. and Hauschke, D. 1999. 'Approximate Sample Sizes for Testing Hypotheses about the Ratio and Difference of Two Means.' Journal of Biopharmaceutical Studies, Volume 9, No. 4, pages 641-650.

Hauschke, D., Kieser, M., Diletti, E., Burke, M. 1999. 'Sample Size Determination for Proving Equivalence Based on the Ratio of Two Means for Normally Distributed Data.' Statistics in Medicine, Volume 18, pages 93-105.

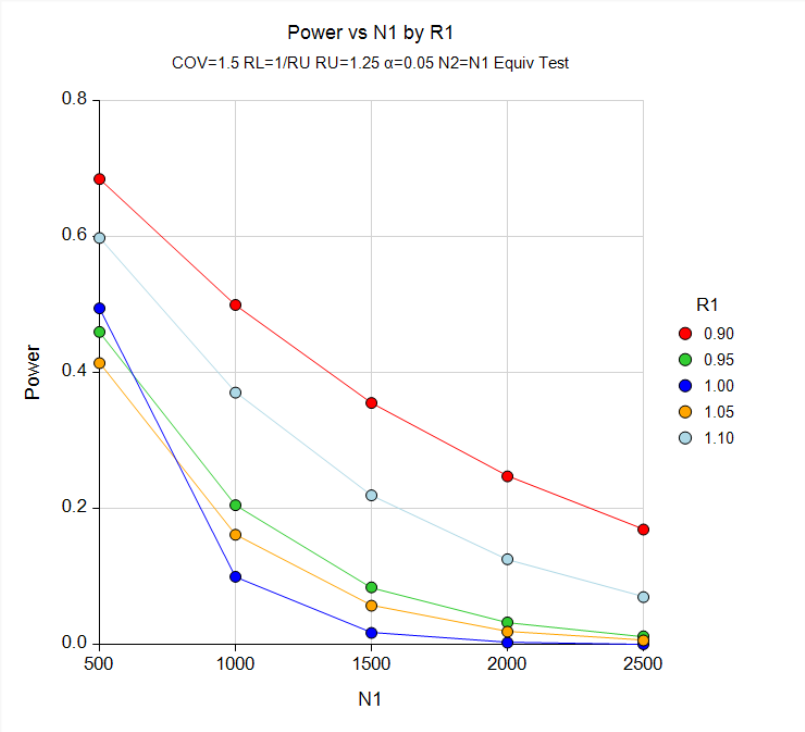
Hauschke, D., Steinijans, V., Pigeot, I. 2007. Bioequivalence Studies in Drug Development. John Wiley and Sons. New York.

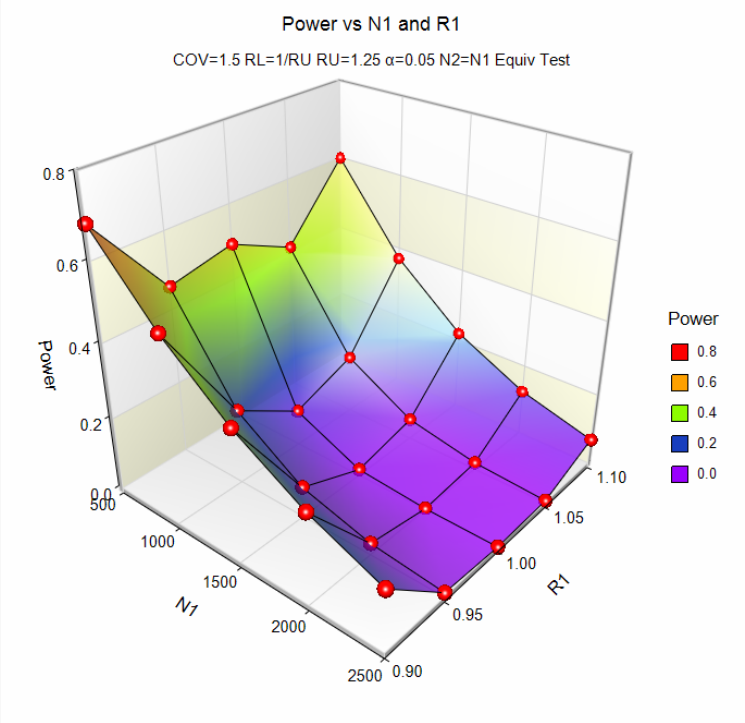
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.

This report shows the power for the indicated scenarios.

Plots Section

Plots





These plots show the power versus the sample size for values of R1.

Example 2 – Validation using Kieser and Hauschke (1999)

Kieser and Hauschke (1999) page 648 present a table of sample sizes for various scenarios. We will use one example to validate this procedure. Let the actual ratio be 1.0, the coefficient of variation be 0.35, the equivalence limits be 0.80 and 1.25, the power be 80%, and the significance level is 0.05. The calculated group sample size is 44.

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	**0.8**
Alpha	**0.05**
Equivalence Limit Input Type	**Enter Upper Equivalence Limit RU (RL = 1 / RU)**
RU (Upper Equivalence Limit)	**1.25**
R1 (Actual Ratio)	**1**
COV (Coefficient of Variation)	**0.35**

Output

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

Numeric Results

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

Power		Sample Size			Mean Ratio				
					Equivalence Limits				
Target	Actual	N1	N2	N	Lower RL	Upper RU	Actual R1	Coefficient of Variation COV	Alpha
0.8	0.80349	44	44	88	0.8	1.25	1	0.35	0.05

PASS also calculates the per group sample size to be 44, which validates the procedure.