Chapter 593

Equivalence Tests for the Ratio of Two Means in a 2x2 Cross-Over Design (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 mixed model. This routine deals with the case in which the statistical hypotheses are expressed in terms of mean ratios rather than mean differences. Also, it assumes that the data follow the normal distribution, so the data will not be logged.

Equivalence Testing Using Ratios of Means

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_C	Not used	*Control (Reference) mean*. This is the mean of a control population.
θ_L, θ_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.
R	R1	*Actual ratio*. This is the value of $R = \mu_T/\mu_C$ at which the power is calculated.

Note that the actual values of μ_T and μ_C 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

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

The alternative hypothesis of equivalence is

$$H_1: R_L < R < R_U.$$

Technical Details

PASS follows the *two one-sided tests* approach described by Schuirmann (1987) and Phillips (1990). The power calculations used in this procedure use the extensions for the mean ratios that were proposed in Hauschke, Kieser, Diletti, and Burke (1999). This procedure also uses the approximations given in Kieser and Hauschke (1999).

Test Statistics

A mixed model may be used to analyze the data from a 2 by 2 cross-over design. This design assumes that subjects are randomized to one of two possible sequences. Those in sequence 1 receive treatment first, then a washout period occurs, and finally the control regimen is given. Those in sequence 2 receive the control first, and then the treatment.

Let Y_{ijk} denote the response of the *j*th subject in the *i*th sequence during period *k.* The following bivariate model is used during the analysis.

$$Y_{ijk} = \mu_h + \pi_k + \varepsilon_{ijk}, \qquad i,k = 1,2 \text{ and } j = 1, \ldots, N_I$$

where μ_h is the effect of treatment *h*, $h = T$ if $i = k$ and $h = C$ if $i \neq k$, π_k is the effect of the *k*th period, $\pi_1 + \pi_2 = 0$. The model assumes that the bivariate vectors $(\varepsilon_{ij1}, \varepsilon_{ij2})'$ are mutually independent and distributed as a bivariate normal with means 0 and covariance matrices Σ_i, where

$$\Sigma_1 = \begin{pmatrix} \sigma_T^2 & \sigma_{TC} \\ \sigma_{TC} & \sigma_C^2 \end{pmatrix} \text{ and } \Sigma_2 = \begin{pmatrix} \sigma_C^2 & \sigma_{TC} \\ \sigma_{TC} & \sigma_T^2 \end{pmatrix}$$

denote the non-singular variance-covariance matrices.

The size-α likelihood ratio test rejects H_{01} and H_{02} if and only if

$$T_1 \geq t_{\alpha,2n-2} \text{ and } T_2 \leq -t_{\alpha,2n-2}$$

where

$$T_i = \frac{\bar{Y}_T - \theta_i \bar{Y}_C}{S_i^2 \sqrt{\frac{1}{4}\left(\frac{2}{N_i}\right)}}$$

$$\bar{Y}_T = \frac{\bar{Y}_{1.1} - \bar{Y}_{2.2}}{2}$$

$$\bar{Y}_C = \frac{\bar{Y}_{1.2} - \bar{Y}_{2.1}}{2}$$

$$S_i^2 = S_T^2 - 2\theta_i S_{TC} + \theta_i^2 S_C^2$$

Further details on computing the test statistics are given in Hauschke et al. (1999).

Power Calculation

Restricting the general bivariate model above to the mixed model yields

$$(\varepsilon_{ij1}, \varepsilon_{ij2})' = (b_{ij} + e_{ij1}, b_{ij} + e_{ij2})'$$

where b_{ij} is the effect of the *j*th subject in the *i*th sequence and e_{ijk} is the within-subject error. These variables are mutually independent with population means 0 and variances σ_b^2 and σ_e^2, respectively. The associated standard deviations can be expressed as proportions of the control mean to form two coefficients of variation as follows.

$$CV_b = \sigma_b/\mu_C$$

$$CV_e = \sigma_e/\mu_C$$

The power is given by

$$\text{Power} = \Pr\left[T_1 \geq t_{\alpha,2n-2} \text{ and } T_2 \leq -t_{\alpha,2n-2} | \theta_1 < \frac{\mu_T}{\mu_C} < \theta_2, \sigma_i^2\right]$$

where $\sigma_i^2 = \sigma_e^2(1 + \theta_i^2) + \sigma_b^2(1 - \theta_i)^2, i = 1,2.$

These test statistics have a bivariate non-central *t*- distribution with non-centrality parameters given by

$$\lambda_i = \frac{\frac{\mu_T}{\mu_C} - \theta_i}{\sqrt{\frac{CV_e^2(1 + \theta_i^2) + CV_b^2(1 - \theta_i)^2}{2N_I}}}, i = 1,2$$

PASS uses the approximation given in Kieser and Hauschke (1999) to calculate this probability. In comparing this approximation to the exact tables in Hauschke et al. (1999), we found that the sample sizes matched exactly except when R1 was 0.95 or 1.05 and one case when it was 0.85. In these cases, the sample size computed by the approximation was one less than that computed by the exact formula.

Obtaining Estimates of CV_b^2 and CV_e^2

It may be difficult to obtain estimates of CV_b and CV_e. This section will show how to calculate the necessary values from a complete example given in Chow and Liu (2009, Chapter 3).

Chapter 3 presents a complete example of how to analyze a standard 2x2 cross-over design using analysis of variance. On page 71, Chow and Liu provide the data for their example. They analyze the data using the following AOV model

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{(j,k)} + +C_{(j-1,k)} + e_{ijk}$$

For this model and data, on page 75 Chow and Liu present the following AOV table which we summarize here.

Source	df	Sum of Squares	Mean Squares
Between			
Carryover	1	276.00	276.00
Residuals	22	16211.49	**736.89**
Within			
Direct Drug	1	62.79	62.79
Period	1	35.97	35.97
Residuals	22	3679.43	**167.25**
Total	47	20265.68	

The estimated value of the control mean is 82.56.

Using the numbers from this example, the necessary constants can be estimated as follows:

$$\widehat{CV_e} = \frac{\sqrt{167.25}}{82.56} = 0.156644$$

$$\widehat{CV_b} = \frac{\sqrt{\frac{(736.89 - 167.25)}{2}}}{82.56} = 0.204416$$

Example 1 – Finding Power

A cross-over design will be used to test the equivalence of two drugs. Researchers have decided to set the equivalence limits for the ratio at 0.80 and 1.25. Past experience leads the researchers to set CVb = 0.8 and CVe = 0.6, 0.7, 0.8 . The significance level is 0.05. The power will be computed assuming that the true ratio is one. Sample sizes between 40 and 160 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**
Sample Size Per Sequence	**40 80 120 160**
Equivalence Limit Input Type	**Enter Upper Equivalence Limit RU (RL = 1 / RU)**
RU (Upper Equivalence Limit)	**1.25**
R1 (Actual Ratio)	**1.0**
CVb (Between Subjects Coef of Variation)	**0.8**
CVe (Within-Subject Coef of Variation)	**0.6 0.7 0.8**

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 / Control Mean
Hypotheses: H0: R ≤ RL or R ≥ RU vs. H1: RL < R < RU

	Sample Size		Mean Ratio			Coefficient of Variation		
			Equivalence Limits					
Power	Sequence Ni	Total N	Lower RL	Upper RU	Actual R1	Between-Subjects CVb	Within-Subject Cve	Alpha
0.46113	40	80	0.8	1.25	1	0.8	0.6	0.05
0.23775	40	80	0.8	1.25	1	0.8	0.7	0.05
0.05208	40	80	0.8	1.25	1	0.8	0.8	0.05
0.88232	80	160	0.8	1.25	1	0.8	0.6	0.05
0.73788	80	160	0.8	1.25	1	0.8	0.7	0.05
0.56721	80	160	0.8	1.25	1	0.8	0.8	0.05
0.97807	120	240	0.8	1.25	1	0.8	0.6	0.05
0.91954	120	240	0.8	1.25	1	0.8	0.7	0.05
0.81803	120	240	0.8	1.25	1	0.8	0.8	0.05
0.99631	160	320	0.8	1.25	1	0.8	0.6	0.05
0.97714	160	320	0.8	1.25	1	0.8	0.7	0.05
0.92784	160	320	0.8	1.25	1	0.8	0.8	0.05

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
Ni	The sample size of each sequence. It is also the number of subjects.
N	The total sample size. N = 2 × Ni.
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.
CVb	The between-subjects coefficient of variation on the original scale. CVb = σb / μC.
CVe	The within-subject coefficient of variation on the original scale. CVe = σe / μC.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A 2×2 cross-over design will be used to test whether the treatment mean (μT) is equivalent to the control mean (μC), with mean ratio equivalence limits of 0.8 and 1.25 (H0: R ≤ 0.8 or R ≥ 1.25 versus H1: 0.8 < R < 1.25, R = μT / μC). The comparison will be made using two one-sided mixed-model-analysis t-tests with an overall Type I error rate (α) of 0.05. This test assumes that the original (untransformed) data in each group follow a normal distribution with a common variance for both groups. The between-subjects coefficient of variation on the original scale is assumed to be 0.8, and the within-subject coefficient of variation on the original scale is assumed to be 0.6. To detect a ratio of means (μT / μC) of 1 with 40 subjects per sequence (for a total of 80 subjects), the power is 0.46113.

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%	40	40	80	50	50	100	10	10	20
20%	80	80	160	100	100	200	20	20	40
20%	120	120	240	150	150	300	30	30	60
20%	160	160	320	200	200	400	40	40	80

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, 50 subjects should be enrolled in Group 1, and 50 in Group 2, to obtain final group sample sizes of 40 and 40, respectively.

References

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.

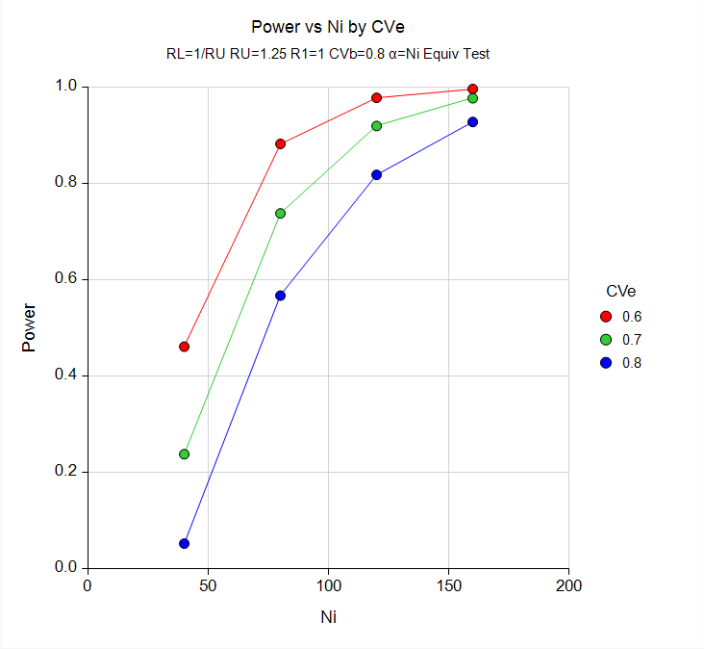
Kieser, M. and Hauschke, D. 1999. 'Approximate Sample Sizes for Testing Hypotheses about the Ratio and Difference of Two Means.' Journal of Biopharmaceutical Studies, Volume 9, No. 4, pages 641-650.

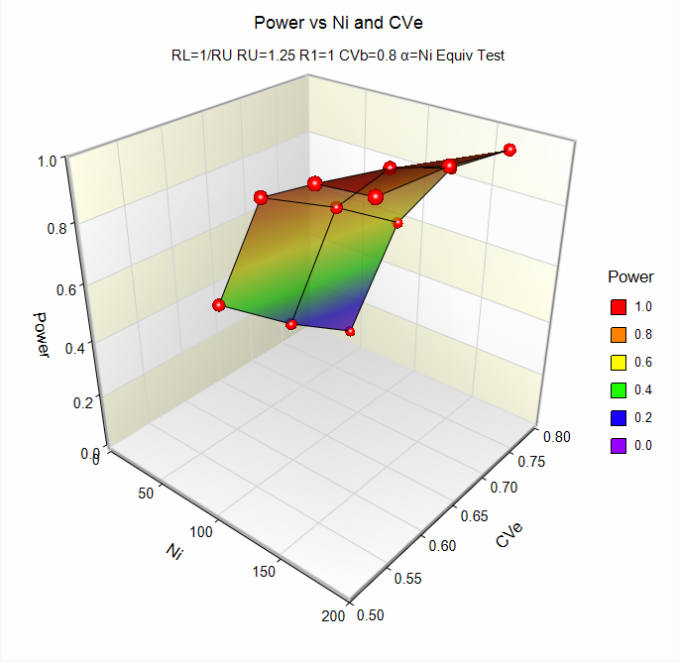
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.

This report shows the power for the indicated scenarios.

Plots Section

Plots





These plots show the relationship between power, sample size, and CVe.

Example 2 – Validation using Hauschke et al. (1999)

Hauschke et al. (1999) page 101 presents a table of sample sizes for various parameter values. When the target power is 0.8, the significance level is 0.05, CVb is 0.2, CVe is 0.3, RL is 0.8, RU is 1.25, and R1 is 1.0, the required sample size per sequence is 17.

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)	**1.0**
CVb (Between Subjects Coef of Variation)	**0.2**
CVe (Within-Subject Coef of Variation)	**0.3**

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			Coefficient of Variation		
				Equivalence Limits					
Target	Actual	Sequence Ni	Total N	Lower RL	Upper RU	Actual R1	Between-Subjects CVb	Within-Subject Cve	Alpha
0.8	0.80964	17	34	0.8	1.25	1	0.2	0.3	0.05

PASS also obtains a sequence size of 17. Thus, the procedure is validated.