Chapter 144

Equivalence Tests for the Ratio of Two Within-Subject Variances in a 2×2M Replicated Cross-Over Design

Introduction

This procedure calculates power and sample size of *equivalence* (or *biosimilarity*) tests of within-subject variabilities from a 2×2M replicated cross-over design. This routine deals with the case in which the statistical hypotheses are expressed in terms of the ratio of the within-subject variances.

This design is used to compare two treatments which are administered to subjects in different orders. It has two treatment sequences. Here, *M* is the number of times a particular treatment is received by a subject. For example, if *M* = 2, the design is a 2×4 cross-over. The two sequences would often be

sequence 1: R T R T

sequence 2: T R T R

It is assumed that either there is no carry-over from one measurement to the next, or there is an ample washout period between measurements.

Technical Details

This procedure uses the formulation given in Chow, Shao, Wang, and Lokhnygina (2018).

Suppose x_{ijkl} is the response in the *i*th sequence (*i* = 1, 2), *j*th subject (*j* = 1, ..., *Ni*), *k*th treatment (*k* = T, C), and *l*th replicate (l = 1, ..., M). The mixed effect model analyzed in this procedure is

$$x_{ijkl} = \mu_k + \gamma_{ikl} + S_{ijk} + e_{ijkl}$$

where μ_k is the *k*th treatment effect, γ_{ikl} is the fixed effect of the *l*th replicate on treatment *k* in the *i*th sequence, S_{ij1} and S_{ij2} are random effects of the *ij*th subject, and e_{ijkl} is the within-subject error term which is normally distributed with mean 0 and variance $V_k = \sigma_{Wk}^2$.

Unbiased estimators of these variances are found after applying an orthogonal transformation matrix *P* to the x's as follows

$$z_{ijk} = P' x_{ijk}$$

where *P* is an *m* × *m* matrix such that $P'P$ is diagonal and $\text{var}(z_{ijkl}) = \sigma_{Wk}^2$.

For example, in a 2×4 cross-over design the z's become

$$z_{ijk1} = \frac{x_{ijk1} + x_{ijk2}}{2} = \bar{x}_{ijk}.$$

and

$$z_{ijk2} = \frac{x_{ijk1} + x_{ijk2}}{\sqrt{2}} = \bar{x}_{ijk}.$$

In this case, the within-subject variances are estimated as

$$\hat{V}_T = \frac{1}{(N_1 + N_2 - 2)(M-1)} \sum_{i=1}^{2} \sum_{j=1}^{N_i} \sum_{l=1}^{M} \left(z_{ijTl} - \bar{z}_{i.Tl}\right)^2$$

and

$$\hat{V}_C = \frac{1}{(N_1 + N_2 - 2)(M-1)} \sum_{i=1}^{2} \sum_{j=1}^{N_i} \sum_{l=1}^{M} \left(z_{ijCl} - \bar{z}_{i.Cl}\right)^2$$

Testing Equivalence

The following hypotheses are usually used to test for equivalence

$$H_0: \frac{\sigma_{WT}^2}{\sigma_{WC}^2} \geq RU \text{ or } \frac{\sigma_{WT}^2}{\sigma_{WC}^2} \leq RL \quad \text{versus} \quad H_1: RL < \frac{\sigma_{WT}^2}{\sigma_{WC}^2} < RU,$$

where RL and RU are the equivalence limits.

These hypotheses can be tested using the two one-sided hypotheses

$$H_{01}: \frac{\sigma_{WT}^2}{\sigma_{WC}^2} \geq RU \quad \text{versus} \quad H_{11}: \frac{\sigma_{WT}^2}{\sigma_{WC}^2} < RU$$

and

$$H_{02}: \frac{\sigma_{WT}^2}{\sigma_{WC}^2} \leq RL \quad \text{versus} \quad H_{12}: \frac{\sigma_{WT}^2}{\sigma_{WC}^2} > RL$$

The corresponding test statistics are $T_1 = RU\ (\hat{V}_1/\hat{V}_2)$ and $T_2 = RL\ (\hat{V}_1/\hat{V}_2)$. Upon making the usual normality assumptions, these *T's* are distributed as $F_{d,d}$ random variables where

$$d = (N_1 + N_2 - 2)(M-1).$$

Power

The power of this combination of tests is given by

$$\text{Power} = \text{P}\left(\frac{RL}{R1}\,F_{1-\alpha,d,d} < F < \frac{RU}{R1}\,F_{\alpha,d,d}\right)$$

where *F* is the common F distribution with the indicated degrees of freedom, α is the significance level, and *R1* is the value of the variance ratio stated by the alternative hypothesis. Lower quantiles of F are used in the equation.

A simple binary search algorithm can be applied to this power function to obtain an estimate of the necessary sample size.

Example 1 – Finding Sample Size

A company has developed a generic drug for treating rheumatism and wants to show that it is equivalent to the standard drug with respect to the within-subject variance. A 2 x 4 cross-over design will be used to test the equivalence of the two drugs.

Company researchers set the upper limit of equivalence to 1.5, the lower limit to 1/1.5, the significance level to 0.05, the power to 0.90, M to 2, and the actual variance ratio values between 0.8 and 1.3. They want to investigate the range of required sample size values assuming that the two group sample sizes are equal.

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	**Sample Size**
Power	**0.90**
Alpha	**0.05**
Sequence Allocation	**Equal (N1 = N2)**
M (Number of Replicates)	**2**
RU (Upper Equivalence Limit)	**1.5**
RL (Lower Equivalence Limit)	**1/RU**
R1 (Actual Variance Ratio)	**0.8 0.9 1 1.1 1.2 1.3**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Variance Ratio: $\sigma^2_{WT} / \sigma^2_{WC}$
Hypotheses: H0: $\sigma^2_{WT} / \sigma^2_{WC} \le RL$ or $\sigma^2_{WT} / \sigma^2_{WC} \ge RU$ vs. H1: $RL < \sigma^2_{WT} / \sigma^2_{WC} < RU$

Power		Sequence Sample Size			Number of Replicates	Variance Ratio: Equivalence Limits			
Target	Actual	N1	N2	N	M	Lower RL	Upper RU	Actual R1	Alpha
0.9	0.9002	517	517	1034	2	0.667	1.5	0.8	0.05
0.9	0.9001	192	192	384	2	0.667	1.5	0.9	0.05
0.9	0.9022	134	134	268	2	0.667	1.5	1.0	0.05
0.9	0.9012	181	181	362	2	0.667	1.5	1.1	0.05
0.9	0.9004	346	346	692	2	0.667	1.5	1.2	0.05
0.9	0.9000	838	838	1676	2	0.667	1.5	1.3	0.05

Target Power: The desired power value entered in the procedure. Power is the probability of rejecting a false null hypothesis.
Actual Power: The actual power achieved. Because N1 and N2 are discrete, this value is usually slightly larger than the target power.
N1: The number of subjects from sequence 1.
N2: The number of subjects from sequence 2.
N: The total number of subjects. N = N1 + N2.
M: The number of replicates. That is, it is the number of times a treatment measurement is repeated on a subject.
RL: The lower equivalence (similarity) limit for the within-subject variance ratio.
RU: The upper equivalence (similarity) limit for the within-subject variance ratio.
R1: The value of the within-subject variance ratio at which the power is calculated.
Alpha: The probability of rejecting a true null hypothesis.

Summary Statements

A 2×2M replicated cross-over design will be used to test whether the treatment within-subject variance (σ^2_{WT}) is equivalent to the control within-subject variance (σ^2_{WC}), by testing whether the within-subject variance ratio ($\sigma^2_{WT} / \sigma^2_{WC}$) is between 0.667 and 1.5 (H0: $\sigma^2_{WT} / \sigma^2_{WC} \le 0.667$ or $\sigma^2_{WT} / \sigma^2_{WC} \ge 1.5$ versus H1: $0.667 < \sigma^2_{WT} / \sigma^2_{WC} < 1.5$). Each subject will alternate treatments (T and C), with an assumed wash-out period between measurements to avoid carry-over. With 2 replicate pairs, each subject will be measured 4 times. For those in the Sequence 1 group, the first treatment will be C, and the sequence is [C T C T]. For those in the Sequence 2 group, the first treatment will be T, and the sequence is [T C T C]. The comparison will be made using two one-sided, variance-ratio F-tests (with the treatment within-subject variance in the numerator), with an overall Type I error rate (α) of 0.05. To detect a within-subject variance ratio ($\sigma^2_{WT} / \sigma^2_{WC}$) of 0.8 with 90% power, the number of subjects needed will be 517 in Group/Sequence 1, and 517 in Group/Sequence 2.

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%	517	517	1034	647	647	1294	130	130	260
20%	192	192	384	240	240	480	48	48	96
20%	134	134	268	168	168	336	34	34	68
20%	181	181	362	227	227	454	46	46	92
20%	346	346	692	433	433	866	87	87	174
20%	838	838	1676	1048	1048	2096	210	210	420

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. 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. After solving for N1 and N2, 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, 647 subjects should be enrolled in Group 1, and 647 in Group 2, to obtain final group sample sizes of 517 and 517, respectively.

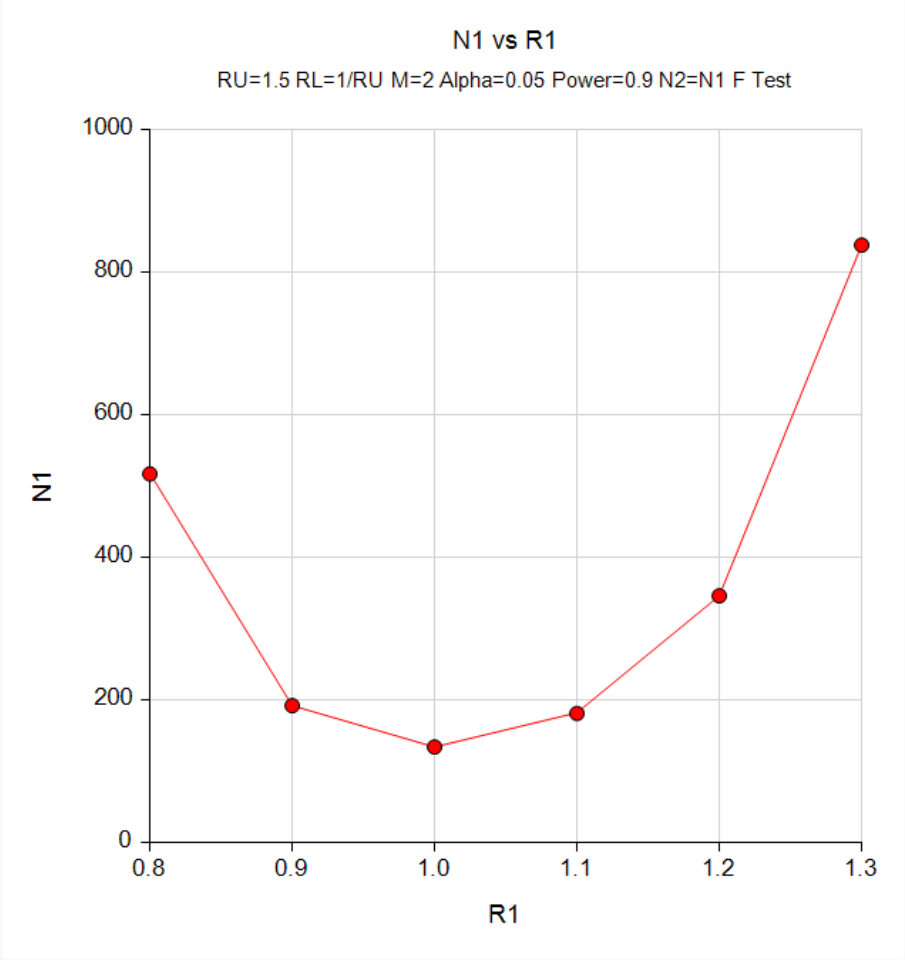
References

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.
Chow, S.C., and Liu, J.P. 2014. Design and Analysis of Clinical Trials, Third Edition. John Wiley & Sons. Hoboken, New Jersey.
Chow, S.C. 2014. Biosimilars Design and Analysis of Follow-on Biologics, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.

This report gives the sample sizes for the indicated scenarios.

Plots Section

Plots



This plot shows the relationship between sample size and R1.

Example 2 – Validation using Hand Calculations

We could not find an example in the literature, so we will present hand calculations to validate this procedure.

Set N1 to 100, the upper limit of equivalence to 1.5, the lower limit to 1/1.5, the significance level to 0.05, M to 2, and the actual variance ratio values 1.0. Compute the power.

The calculations proceed as follows.

$$\begin{aligned}\text{Power} &= \text{P}\left(\frac{RL}{R1}\,F_{1-\alpha,d,d} < F < \frac{RU}{R1}\,F_{\alpha,d,d}\right)\\ &= \text{P}\left(0.666667/1\left(F_{0.95,198,198}\right) < F < 1.5/1\left(F_{0.05,198,198}\right)\right)\\ &= \text{P}\left(0.666667(1.26408895) < F < 1.5(0.79108357)\right)\\ &= \text{P}(0.84272639 < F < 1.18662536)\\ &= 0.88524837 - 0.11475232\\ &= 0.77049605\end{aligned}$$

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	**Power**
Alpha	**0.05**
Sequence Allocation	**Equal (N1 = N2)**
Sample Size Per Sequence	**100**
M (Number of Replicates)	**2**
RU (Upper Equivalence Limit)	**1.5**
RL (Lower Equivalence Limit)	**1/RU**
R1 (Actual Variance Ratio)	**1.0**

Output

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

Numeric Results

Solve For: Power
Variance Ratio: $\sigma^2_{WT} / \sigma^2_{WC}$
Hypotheses: H0: $\sigma^2_{WT} / \sigma^2_{WC} \le RL$ or $\sigma^2_{WT} / \sigma^2_{WC} \ge RU$ vs. H1: $RL < \sigma^2_{WT} / \sigma^2_{WC} < RU$

					Variance Ratio			
	Sequence Sample Size			Number of Replicates	Equivalence Limits			
Power	N1	N2	N	M	Lower RL	Upper RU	Actual R1	Alpha
0.7705	100	100	200	2	0.667	1.5	1	0.05

The power matches the hand-calculated result.