Chapter 317

Non-Unity Null Tests for Two Between Variances in a Replicated Design

Introduction

This procedure calculates power and sample size of tests of the between-subject variance (between + within) from a parallel (two-group) design with replicates (repeated measures) for the case when the ratio assumed by the null hypothesis is not necessarily one. This routine deals with the case in which the statistical hypotheses are expressed in terms of the ratio of the between-subject variances.

A parallel design is used to compare two treatment groups by comparing subjects receiving each treatment. In this replicated design, each subject is measured *M* times where *M* is at least two. To be clear, each subject receives only one treatment, but is measured repeatedly.

Replicated parallel designs such as this are popular because they allow the assessment of total variances, between-subject variances, and within-subject variances.

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), pages 209 - 212.

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

$$x_{ijk} = \mu_i + S_{ij} + e_{ijk}$$

where μ_i is the treatment effect, S_{ij} is the random effect of the *j*th subject in the *i*th treatment, and e_{ijk} is the within-subject error term which is normally distributed with mean 0 and variance $V_i = \sigma_{Wi}^2$.

Unbiased estimates of these variances are given by

$$\hat{\sigma}_{Wi}^2 = s_{Wi}^2 = \frac{1}{N_i(M-1)} \sum_{j=1}^{N_i} \sum_{k=1}^{M} \left(x_{ijk} - \bar{x}_{ij.}\right)^2 \text{ , } i = T, C$$

where

$$\bar{x}_{ij.} = \frac{1}{M} \sum_{k=1}^{M} x_{ijk}$$

Define

$$s_{Bi}^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(\bar{x}_{ij.} - \bar{x}_{i..}\right)^2$$

where

$$\bar{x}_{i..} = \frac{1}{N_i}\sum_{j=1}^{N_i}\bar{x}_{ij.}$$

Now, estimators for the between-subject variance are given by

$$\hat{\sigma}_{Bi}^2 = s_{Bi}^2 - \frac{1}{M}\hat{\sigma}_{Wi}^2$$

Testing Variance Inequality with a Non-Unity Null

The following three sets of statistical hypotheses are used to test for between-subject variance inequality with a non-unity null

$$H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \geq R0 \quad \text{versus} \quad H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} < R0,$$

$$H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \leq R0 \quad \text{versus} \quad H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} > R0,$$

$$H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} = R0 \quad \text{versus} \quad H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \neq R0,$$

where *R0* is the variance ratio assumed by the null hypothesis (usually, one).

Let $\eta = \sigma_{BT}^2 - R0(\sigma_{BC}^2)$ be the parameter of interest. The test statistic is $\hat{\eta} = \hat{\sigma}_{BT}^2 - R0(\hat{\sigma}_{BC}^2)$.

Two-Sided Test

For the two-sided test, compute two limits, $\hat{\eta}_L$ and $\hat{\eta}_U$, using

$$\hat{\eta}_L = \hat{\eta} - \sqrt{\Delta_L}$$

$$\hat{\eta}_U = \hat{\eta} + \sqrt{\Delta_U}$$

Reject the null hypothesis if $\hat{\eta}_L > 0$ is or $\hat{\eta}_U < 0$.

The $\Delta's$ are given by

$$\Delta_L = h\left(\frac{\alpha}{2}, N_T - 1\right) s_{BT}^4 + h\left(1 - \frac{\alpha}{2}, N_C - 1\right) R_0^2 s_{BC}^4 + h\left(1 - \frac{\alpha}{2}, N_T(M-1)\right)\left[\frac{s_{WT}^2}{M}\right]^2 + h\left(\frac{\alpha}{2}, N_C(M-1)\right)\left[\frac{R_0 s_{WC}^2}{M}\right]^2$$

$$\Delta_U = h\left(1 - \frac{\alpha}{2}, N_T - 1\right) s_{BT}^4 + h\left(\frac{\alpha}{2}, N_C - 1\right) R_0^2 s_{BC}^4 + h\left(\frac{\alpha}{2}, N_T(M-1)\right)\left[\frac{s_{WT}^2}{M}\right]^2 + h\left(1 - \frac{\alpha}{2}, N_C(M-1)\right)\left[\frac{R_0 s_{WC}^2}{M}\right]^2$$

where

$$h(A, B) = \left(1 - \frac{B}{\chi_{A,B}^2}\right)^2$$

and $\chi_{A,B}^2$ is the upper quantile of the chi-square distribution with B degrees of freedom.

One-Sided Test

For the lower, one-sided test, compute the limit, $\hat{\eta}_U$, using

$$\hat{\eta}_U = \hat{\eta} + \sqrt{\Delta_U}$$

Reject the null hypothesis if $\hat{\eta}_U < 0$.

The Δ_U is given by

$$\Delta_U = h(1 - \alpha, N_T - 1) s_{BT}^4 + h(\alpha, N_C - 1) R_0^2 s_{BC}^4 + h\big(\alpha, N_T(M-1)\big)\left[\frac{s_{WT}^2}{M}\right]^2 + h\big(1 - \alpha, N_C(M-1)\big)\left[\frac{R_0 s_{WC}^2}{M}\right]^2$$

Power

Two-Sided Test

The power of the two-sided test assuming $n = N_T = N_C$ is given by

$$\text{Power} = 1 - \Phi\left(z_{1-\frac{\alpha}{2}} - \frac{(R_1 - R_0)\sigma_{BC}^2}{\sqrt{\sigma^{*2}/n}}\right) + \Phi\left(z_{\alpha/2} - \frac{(R_1 - R_0)\sigma_{BC}^2}{\sqrt{\sigma^{*2}/n}}\right)$$

where

$$R_1 = \frac{\sigma_{BT}^2}{\sigma_{BC}^2}$$

$$\sigma_{BT}^2 = R_1 \sigma_{BC}^2$$

$$\sigma^{*2} = 2\left[\left(\sigma_{BT}^2 + \frac{\sigma_{WT}^2}{M}\right)^2 + R_0^2\left(\sigma_{BC}^2 + \frac{\sigma_{WC}^2}{M}\right)^2 + \frac{\sigma_{WT}^4}{M^2(M-1)} + \frac{R_0^2\sigma_{WC}^4}{M^2(M-1)}\right]$$

where *R1* is the value of the variance ratio stated by the alternative hypothesis and $\Phi(x)$ is the standard normal CDF.

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

One-Sided Test

The power of the lower, one-sided test, $H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \geq R0$ versus $H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} < R0$, is given by

$$\text{Power} = \Phi\left(z_{\alpha} - \frac{(R_1 - R_0)\sigma_{BC}^2}{\sqrt{\sigma^{*2}/n}}\right)$$

The power of the upper, one-sided test, $H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \leq R0$ versus $H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} > R0$, is given by

$$\text{Power} = 1 - \Phi\left(z_{1-\alpha} - \frac{(R_1 - R_0)\sigma_{BC}^2}{\sqrt{\sigma^{*2}/n}}\right)$$

Example 1 – Finding Sample Size

A company has developed a generic drug for treating rheumatism and wants to compare it to the standard drug in terms of the between-subject variability. A two-group, parallel design with replicates will be used to test the inequality using a two-sided test.

Company researchers set the variance ratio under the null hypothesis to 0.8, the significance level to 0.05, the power to 0.90, M to 2, and the actual variance ratio values between 0.5 and 1.3. They also set σ^2_{BC} = 0.8, σ^2_{WT} = 0.2, and σ^2_{WC} = 0.3. 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**
Alternative Hypothesis	**Two-Sided (H1: $\sigma^2_{BT}/\sigma^2_{BC}$ ≠ R0)**
Power	**0.90**
Alpha	**0.05**
M (Measurements Per Subject)	**2**
R0 (H0 Variance Ratio)	**0.8**
R1 (Actual Variance Ratio)	**0.5 0.7 0.9 1.1 1.3**
σ^2_{BC} (Control Variance)	**0.8**
σ^2_{WT} (Treatment Variance)	**0.2**
σ^2_{WC} (Control Variance)	**0.3**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Hypotheses: H0: $\sigma^2_{BT}/\sigma^2_{BC}$ = R0 vs. H1: $\sigma^2_{BT}/\sigma^2_{BC}$ ≠ R0

Power		Sample Size			Measurements	Between-Subject Variance Ratio		Between-Subject Variance	Within-Subject Variance		
Target	Actual	Treatment N_T	Control N_C	Total N	per Subject M	H0 (Null) R0	Actual R1	Control σ^2_{BC}	Treatment σ^2_{WT}	Control σ^2_{WC}	Alpha
0.9	0.9001	311	311	622	2	0.8	0.5	0.8	0.2	0.3	0.05
0.9	0.9001	3408	3408	6816	2	0.8	0.7	0.8	0.2	0.3	0.05
0.9	0.9000	4185	4185	8370	2	0.8	0.9	0.8	0.2	0.3	0.05
0.9	0.9005	571	571	1142	2	0.8	1.1	0.8	0.2	0.3	0.05
0.9	0.9003	250	250	500	2	0.8	1.3	0.8	0.2	0.3	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 N_T and N_C are discrete, this value is usually slightly larger than the target power.
N_T The number of subjects in the treatment group.
N_C The number of subjects in the control group.
N The total number of subjects. N = N_T + N_C.
M The number of times a subject is measured. It is the number of repeated measurements.
R0 The between-subject variance ratio used to define the null hypothesis, H0.
R1 The value of the between-subject variance ratio at which the power is calculated. R1 = σ^2_{BT} / σ^2_{BC}.
σ^2_{BC} The between-subject variance of measurements in the control group. Note that σ^2_{TC} = σ^2_{BC} + σ^2_{WC}.
σ^2_{WT} The within-subject variance of measurements in the treatment group.
σ^2_{WC} The within-subject variance of measurements in the control group.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group replicated design will be used to test whether the between-subject variance ratio (σ^2_{BT} / σ^2_{BC} = σ^2Between,Treatment / σ^2Between,Control) is different from 0.8 (H0: σ^2_{BT} / σ^2_{BC} = 0.8 versus H1: σ^2_{BT} / σ^2_{BC} ≠ 0.8). The comparison will be made using a two-sided, variance-difference test (treatment minus control) as described in Chow, Shao, Wang, and Lokhnygina (2018), with a Type I error rate (α) of 0.05. Each subject will be measured 2 times. For the control group, the between-subject variance (σ^2_{BC}) is assumed to be 0.8, and the within-subject variance is assumed to be 0.3. The within-subject variance of the treatment group is assumed to be 0.2. To detect a between-subject variance ratio (σ^2_{BT} / σ^2_{BC}) of 0.5 with 90% power, the number of subjects needed will be 311 in the treatment group, and 311 in the control group.

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	N_T	N_C	N	N_T'	N_C'	N'	D_T	D_C	D
20%	311	311	622	389	389	778	78	78	156
20%	3408	3408	6816	4260	4260	8520	852	852	1704
20%	4185	4185	8370	5232	5232	10464	1047	1047	2094
20%	571	571	1142	714	714	1428	143	143	286
20%	250	250	500	313	313	626	63	63	126

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_T, N_C, and N: The evaluable sample sizes at which power is computed. If N_T and N_C subjects are evaluated out of the N_T' and N_C' subjects that are enrolled in the study, the design will achieve the stated power.

N_T', N_C', and N': The number of subjects that should be enrolled in the study in order to obtain N_T, N_C, and N evaluable subjects, based on the assumed dropout rate. After solving for N_T and N_C, N_T' and N_C' are calculated by inflating N_T and N_C using the formulas $N_T' = N_T / (1 - DR)$ and $N_C' = N_C / (1 - DR)$, with N_T' and N_C' 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_T, D_C, and D: The expected number of dropouts. $D_T = N_T' - N_T$, $D_C = N_C' - N_C$, and $D = D_T + D_C$.

Dropout Summary Statements

Anticipating a 20% dropout rate, 389 subjects should be enrolled in Group 1, and 389 in Group 2, to obtain final group sample sizes of 311 and 311, respectively.

References

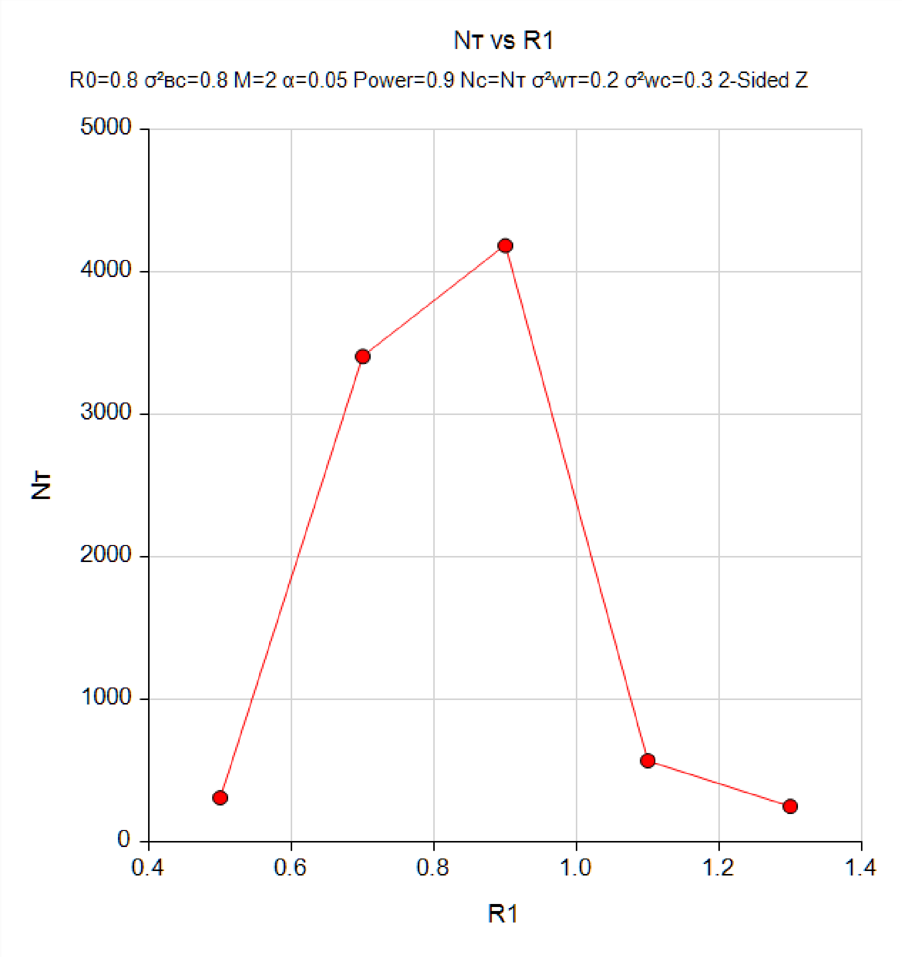
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.

Chow, S.C., and Liu, J.P. 2014. Design and Analysis of Clinical Trials, Third Edition. John Wiley & Sons. Hoboken, New Jersey.

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 Chow et al. (2018)

We will use an example from Chow *et al*. (2018) pages 212-213 to validate this procedure.

In this example, R0 = 1.21, power = 0.8, significance level = 0.05, M = 3, R1 = 0.5625, σ^2_{BC} = 0.16, σ^2_{WT} = 0.04, σ^2_{WC} = 0.09. The problem is to find the sample size for the lower, one-sided test (note that this is a non-inferiority test). They find the per group sample size to be approximately 74.

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**
Alternative Hypothesis .... **One-Sided (H1: $\sigma^2_{BT}/\sigma^2_{BC}$ < R0)**
Power... **0.80**
Alpha.. **0.05**
M (Measurements Per Subject) **3**
R0 (H0 Variance Ratio)................................. **1.21**
R1 (Actual Variance Ratio) **0.5625**
σ^2_{BC} (Control Variance)................................. **0.16**
σ^2_{WT} (Treatment Variance) **0.04**
σ^2_{WC} (Control Variance)................................ **0.09**

Output

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

Numeric Results

Solve For: Sample Size
Hypotheses: H0: $\sigma^2_{BT}/\sigma^2_{TC}$ ≥ R0 vs. H1: $\sigma^2_{BT}/\sigma^2_{BC}$ < R0

Power		Sample Size				Between-Subject Variance Ratio			Within-Subject Variance		
Target	Actual	Treatment N_T	Control N_C	Total N	Measurements per Subject M	H0 (Null) R0	Actual R1	Control σ^2_{BC}	Treatment σ^2_{WT}	Control σ^2_{WC}	Alpha
0.8	0.8044	75	75	150	3	1.21	0.563	0.16	0.04	0.09	0.05

The sample size of 75 per group is close to their answer of 74. The difference occurs because of rounding. The sample size of 74 per group actually has a power slightly less than 0.8.