

## Chapter 318

# Non-Inferiority Tests for Two Between Variances in a Replicated Design

## Introduction

This procedure calculates power and sample size of tests of non-inferiority 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 ( $j = 1, \dots, N_i$ ), and  $k$ th replicate ( $k = 1, \dots, M$ ). The model analyzed in this procedure is

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

where  $\mu_i$  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

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

where

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

## Non-Inferiority Tests for Two Between Variances in a Replicated Design

Define

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

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 Non-Inferiority

The following statistical hypotheses are used to test for between-subject variance non-inferiority.

$$H_0: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} \geq R_0 \quad \text{versus} \quad H_1: \frac{\sigma_{BT}^2}{\sigma_{BC}^2} < R_0,$$

where  $R_0$  is the non-inferiority limit.

Let  $\eta = \sigma_{BT}^2 - R_0(\sigma_{BC}^2)$  be the parameter of interest. The test statistic is  $\hat{\eta} = \hat{\sigma}_{BT}^2 - R_0(\hat{\sigma}_{BC}^2)$ .

### Non-Inferiority Test

For the non-inferiority 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  $\Delta$  is given by

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

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.

## Power

### Non-Inferiority Test

The power of the non-inferiority test is given by

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

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  $R_1$  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.

## Example 1 – Finding Sample Size

A company has developed a generic drug for treating rheumatism and wants to show that it is non-inferior to the standard drug in terms of the between-subject variability. A two-group, parallel design will be used to test the non-inferiority.

Company researchers set the non-inferiority limit to 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 also set  $\sigma^2_{bc} = 0.8$ ,  $\sigma^2_{wt} = 0.2$ , and  $\sigma^2_{wc} = 0.3$ . 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 .....	<b>Sample Size</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
M (Measurements Per Subject) .....	<b>2</b>
R0 (Non-Inferiority Variance Ratio) .....	<b>1.5</b>
R1 (Actual Variance Ratio) .....	<b>0.8 0.9 1 1.1 1.2 1.3</b>
$\sigma^2_{bc}$ (Control Variance).....	<b>0.8</b>
$\sigma^2_{wt}$ (Treatment Variance) .....	<b>0.2</b>
$\sigma^2_{wc}$ (Control Variance).....	<b>0.3</b>

## Output

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

### Numeric Reports

#### Numeric Results

Solve For: [Sample Size](#)  
 Hypotheses:  $H_0: \sigma^2_{BT}/\sigma^2_{BC} \geq R_0$  vs.  $H_1: \sigma^2_{BT}/\sigma^2_{BC} < R_0$

Power		Sample Size			Measurements per Subject M	Between-Subject Variance			Within-Subject Variance		Alpha
						Ratio		Control $\sigma^2_{BC}$	Treatment $\sigma^2_{WT}$	Control $\sigma^2_{WC}$	
Target	Actual	Treatment $N_T$	Control $N_C$	Total N	Non-Inferiority $R_0$	Actual $R_1$	Treatment $\sigma^2_{WT}$				Control $\sigma^2_{WC}$
0.9	0.9015	145	145	290	2	1.5	0.8	0.8	0.2	0.3	0.05
0.9	0.9007	206	206	412	2	1.5	0.9	0.8	0.2	0.3	0.05
0.9	0.9004	311	311	622	2	1.5	1.0	0.8	0.2	0.3	0.05
0.9	0.9003	511	511	1022	2	1.5	1.1	0.8	0.2	0.3	0.05
0.9	0.9000	956	956	1912	2	1.5	1.2	0.8	0.2	0.3	0.05
0.9	0.9000	2269	2269	4538	2	1.5	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.
- $R_0$  The non-inferiority limit for the between-subject variance ratio.
- $R_1$  The value of the between-subject variance ratio at which the power is calculated.  $R_1 = \sigma^2_{BT} / \sigma^2_{BC}$ .
- $\sigma^2_{BC}$  The between-subject variance of measurements in the control group. Note that  $\sigma^2_{TC} = \sigma^2_{BC} + \sigma^2_{WC}$ .
- $\sigma^2_{WT}$  The within-subject variance of measurements in the treatment group.
- $\sigma^2_{WC}$  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 of the treatment ( $\sigma^2_{BT}$ ) is non-inferior to the between-subject variance of the control ( $\sigma^2_{BC}$ ) by testing whether the between-subject variance ratio ( $\sigma^2_{BT} / \sigma^2_{BC}$ ) is less than the non-inferiority ratio 1.5 ( $H_0: \sigma^2_{BT} / \sigma^2_{BC} \geq 1.5$  versus  $H_1: \sigma^2_{BT} / \sigma^2_{BC} < 1.5$ ). The comparison will be made using a one-sided, variance-difference test (treatment minus control) as described in Chow, Shao, Wang, and Lohknygina (2018), with a Type I error rate ( $\alpha$ ) of 0.05. Each subject will be measured 2 times. For the control group, the between-subject variance ( $\sigma^2_{BC}$ ) 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 ( $\sigma^2_{BT} / \sigma^2_{BC}$ ) of 0.8 with 90% power, the number of subjects needed will be 145 in the treatment group, and 145 in the control group.

## Non-Inferiority Tests for Two Between Variances in a Replicated Design

## Dropout-Inflated Sample Size

Dropout Rate	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
	N <sub>T</sub>	N <sub>C</sub>	N	N <sub>T</sub> '	N <sub>C</sub> '	N'	D <sub>T</sub>	D <sub>C</sub>	D
20%	145	145	290	182	182	364	37	37	74
20%	206	206	412	258	258	516	52	52	104
20%	311	311	622	389	389	778	78	78	156
20%	511	511	1022	639	639	1278	128	128	256
20%	956	956	1912	1195	1195	2390	239	239	478
20%	2269	2269	4538	2837	2837	5674	568	568	1136

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 <sub>T</sub> , N <sub>C</sub> , and N	The evaluable sample sizes at which power is computed. If N <sub>T</sub> and N <sub>C</sub> subjects are evaluated out of the N <sub>T</sub> ' and N <sub>C</sub> ' subjects that are enrolled in the study, the design will achieve the stated power.
N <sub>T</sub> ', N <sub>C</sub> ', and N'	The number of subjects that should be enrolled in the study in order to obtain N <sub>T</sub> , N <sub>C</sub> , and N evaluable subjects, based on the assumed dropout rate. After solving for N <sub>T</sub> and N <sub>C</sub> , N <sub>T</sub> ' and N <sub>C</sub> ' are calculated by inflating N <sub>T</sub> and N <sub>C</sub> using the formulas $N_{T'} = N_T / (1 - DR)$ and $N_{C'} = N_C / (1 - DR)$ , with N <sub>T</sub> ' and N <sub>C</sub> ' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lohknygina, Y. (2018) pages 32-33.)
D <sub>T</sub> , D <sub>C</sub> , 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, 182 subjects should be enrolled in Group 1, and 182 in Group 2, to obtain final group sample sizes of 145 and 145, respectively.

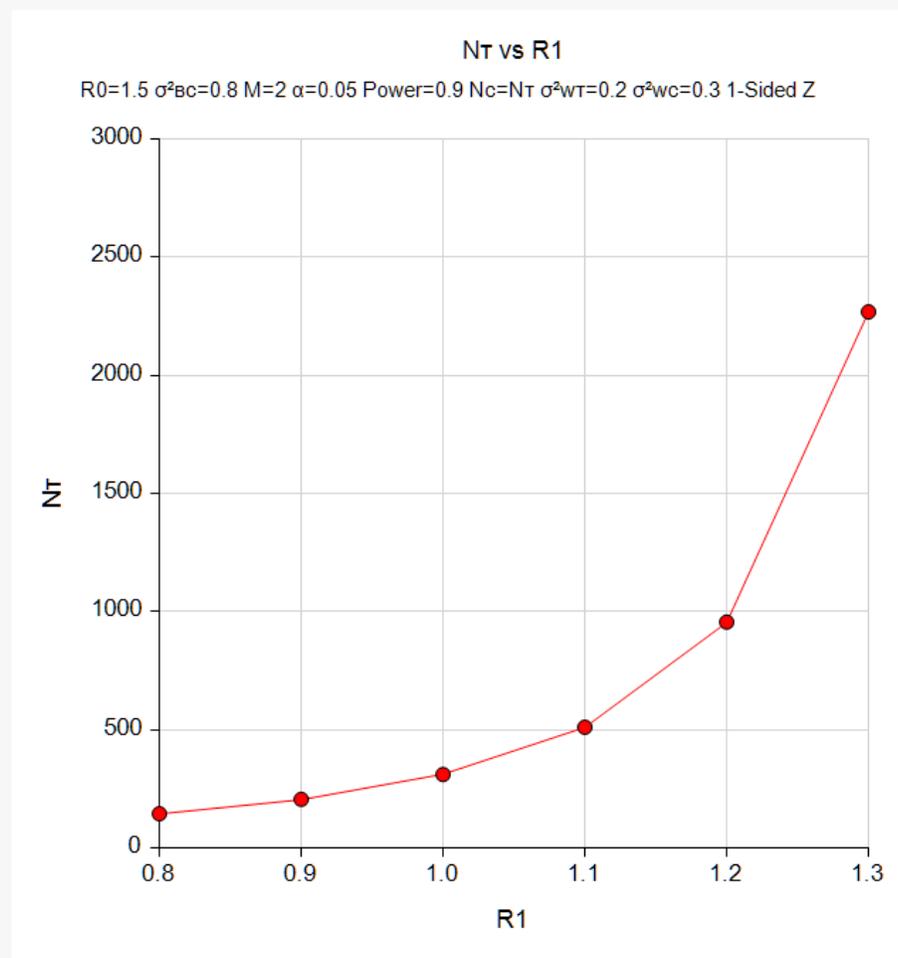
## References

- Chow, S.C., Shao, J., Wang, H., and Lohknygina, 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,  $R_0 = 1.21$ , significance level = 0.05, power = 0.8,  $M = 3$ ,  $R_1 = 0.5625$ ,  $\sigma^2_{bc} = 0.16$ ,  $\sigma^2_{wt} = 0.04$ ,  $\sigma^2_{wc} = 0.09$ . The problem is to find the sample size. 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 .....	<b>Sample Size</b>
Power.....	<b>0.80</b>
Alpha.....	<b>0.05</b>
M (Measurements Per Subject) .....	<b>3</b>
R0 (Non-Inferiority Variance Ratio).....	<b>1.21</b>
R1 (Actual Variance Ratio) .....	<b>0.5625</b>
$\sigma^2_{bc}$ (Control Variance).....	<b>0.16</b>
$\sigma^2_{wt}$ (Treatment Variance) .....	<b>0.04</b>
$\sigma^2_{wc}$ (Control Variance).....	<b>0.09</b>

### Output

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

Numeric Results												
Solve For:		<a href="#">Sample Size</a>										
Hypotheses:		H0: $\sigma^2_{B1}/\sigma^2_{bc} \geq R_0$ vs. H1: $\sigma^2_{B1}/\sigma^2_{bc} < R_0$										
Power		Sample Size			Measurements per Subject M	Between-Subject Variance			Within-Subject Variance			Alpha
Target	Actual	Treatment Nt	Control Nc	Total N		Ratio	Non-Inferiority R0	Actual R1	Control $\sigma^2_{bc}$	Treatment $\sigma^2_{wt}$	Control $\sigma^2_{wc}$	
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.