Chapter 572

Non-Inferiority Tests for the Ratio of Two Means (Normal Data)

Introduction

This procedure calculates power and sample size for *non-inferiority* t-tests from a parallel-groups 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 instead of mean differences.

The details of this analysis are given in Rothmann, Wiens, and Chan (2012) and, to a lesser extent, in Kieser and Hauschke (1999).

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

Non-Inferiority Testing Using Ratios

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 reference population.
R_L	RL	*Lower Non-Inferiority Limit*. This is the lower limit for the mean ratio when higher values are 'better'. Values below this amount are assumed to be inferior. Values above this amount are assumed to be non-inferior.
R_U	RU	*Upper Non-Inferiority Limit*. This is the upper limit for the mean ratio when lower values are 'better'. Values above this amount are assumed to be inferior. Values below this amount are assumed to be non-inferior.
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.

When higher means are better, the hypotheses are arranged so that rejecting the null hypothesis implies that the ratio of the treatment mean to the reference mean is greater than the non-inferiority limit. The value of ϕ at which power is calculated must be greater than $R_L < \phi$.

$$H_0: \phi \leq R_L \quad \text{versus} \quad H_1: \phi > R_L$$

When higher means are worse, the hypotheses are arranged so that rejecting the null hypothesis implies that the ratio of the treatment mean to the reference mean is less than one plus the margin of non-inferiority. The value of ϕ at which power is calculated must be less than $R_U > \phi$.

$$H_0: \phi \geq R_U \quad \text{versus} \quad H_1: \phi < R_U$$

Coefficient of Variation

The coefficient of variation (CV) 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, $CV = \frac{\sigma_C}{\mu_C}$.

Power Calculation

Four tests are provided by Rothmann, Wiens, and Chan (2012) for testing non-inferiority based on the mean ratio when the data are assumed to be normally distributed (untransformed). This section will summarize these tests and the associated power and sample size formulas. Rothmann, Wiens, and Chan (2012) provide a much more complete discussion of this topic.

Tests

This section will provide technical details about the four available test statistics that are available for testing non-inferiority using the mean ratio. We begin with some nomenclature.

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 variances σ_T^2 and σ_C^2. To conduct the comparison, a random sample of N_T and N_C subjects will be obtained for each group. The parameters of the study are presented in terms of the mean ratio $\phi = \mu_T/\mu_C$.

In the results below, let $\lambda = \sigma_T/\sigma_C$, $k = N_T/N_C$, $CV = \sigma_C/\mu_C$, and β be the probability of a type II error.

Assuming that $R_L < 1$, the non-inferiority hypotheses are

$$H_0: \phi \leq R_L \quad \text{versus} \quad H_1: \phi > R_L$$

Four test statistics may be used to test these hypotheses. These are (1) an equal variance t-test, (2) unequal variances large sample z-test, (3) unequal variances Satterthwaite t-test, and (4) unequal variances delta-method z-test.

Equal Variances T-Test

The ratio hypotheses are rearranged as from

$$H_0: \mu_T/\mu_C \leq R_L \quad \text{versus} \quad H_1: \mu_T/\mu_C > R_L$$

to

$$H_0: \mu_T - R_L\mu_C \leq 0 \quad \text{versus} \quad H_1: \mu_T - R_L\mu_C > 0$$

The null hypothesis is tested using the test statistic

$$T_1 = \sqrt{\frac{\bar{X}_T - R_L\bar{X}_C}{S\left(\frac{1}{N_T} + \frac{R_L^2}{N_C}\right)}}$$

where $\bar{X}_T$ and $\bar{X}_C$ are the sample means of the two groups and S is the pooled estimate of the standard deviation, σ which is given by

$$S^2 = \frac{(N_T - 1)s_T^2 + (N_C - 1)s_C^2}{N_T - N_C - 2}$$

It is assumed that T_1 is distributed as a central *t* distribution with degrees of freedom given by $N_T + N_C - 2$.

For a specified alternative $\phi = R_1$, T_1 follows the noncentral *t* distribution with $N_T + N_C - 2$ degrees of freedom and noncentrality

$$\left(\frac{\phi - R_L}{CV}\right)\sqrt{\frac{N_C}{\frac{\lambda^2}{k} + R_L^2}}$$

Hence, the power of this test is given by

$$(1 - \beta) = Pr\left(T_1 \geq t_{1-\alpha, N_T+N_C-2} | \phi, R_L, CV\right)$$

Unequal Variances Large Sample Z-Test

The ratio hypotheses are rearranged as from

$$H_0: \mu_T/\mu_C \leq R_L \quad \text{versus} \quad H_1: \mu_T/\mu_C > R_L$$

to

$$H_0: \mu_T - R_L\mu_C \leq 0 \quad \text{versus} \quad H_1: \mu_T - R_L\mu_C > 0$$

The null hypothesis is tested using the test statistic

$$T_2 = \frac{\bar{x}_T - R_L\bar{x}_C}{\sqrt{\frac{s_T^2}{N_T} + \frac{s_C^2}{N_C}}}$$

where $\bar{x}_T$ and $\bar{x}_C$ are the sample means of the two groups and s_T and s_C are the estimated of the standard deviations.

It is assumed that T_2 has a standard normal distribution when the null hypothesis is true. When $T_2 > z_\alpha$, the null hypothesis is rejected, and non-inferiority is concluded at a one-sided level of α.

Hence, the approximate power of this test is given by

$$z_\beta = \frac{\mu_T - R_L\mu_C}{\sqrt{\frac{\sigma_T^2}{N_T} + R_L^2\frac{\sigma_C^2}{N_C}}} - z_\alpha$$

This can be rearranged to give

$$z_\beta = \sqrt{\frac{N_C}{\frac{\lambda^2}{k} + R_L^2}}\left(\frac{\phi - R_L}{CV}\right) - z_\alpha$$

Unequal Variances Satterthwaite T-Test

The ratio hypotheses are rearranged as from

$$H_0: \mu_T/\mu_C \leq R_L \quad \text{versus} \quad H_1: \mu_T/\mu_C > R_L$$

to

$$H_0: \mu_T - R_L\mu_C \leq 0 \quad \text{versus} \quad H_1: \mu_T - R_L\mu_C > 0$$

The null hypothesis is tested using the test statistic

$$T_3 = \frac{\bar{x}_T - R_L\bar{x}_C}{\sqrt{\frac{\sigma_T^2}{N_T} + R_L^2\frac{\sigma_C^2}{N_C}}}$$

where $\bar{x}_T$ and $\bar{x}_C$ are the sample means of the two groups and s_T and s_C are the estimated of the standard deviations.

It is assumed that the distribution of T_3 is a Satterthwaite adjusted central *t* instead of a standard normal when the null hypothesis is true. When $T_3 > -t_{\alpha,v}$, the null hypothesis is rejected, and non-inferiority is concluded at a one-sided level of α. The Satterthwaite degrees of freedom is given by

$$v = \frac{\left[\frac{s_T^2}{N_T} + R_L^2\frac{s_C^2}{N_C}\right]^2}{\frac{s_T^4}{N_T(N_T-1)} + R_L^4\frac{s_C^4}{N_C(N_C-1)}}$$

The power of this test is given by the non-central *t* distribution with degrees of freedom *v′* estimated by substituting the standard deviations σ_T and σ_C for s_T and s_C in the formula for *v*. The resulting value of *v′* is

$$v' = \frac{\left[\phi^2 + \frac{\lambda^2}{k}\right]^2}{\frac{\lambda^4}{k^2(kN_C-1)} + \frac{R_L^4}{N_C-1}}$$

The non-centrality parameter is given by

$$\left(\frac{\phi - R_L}{CV}\right)\sqrt{\frac{N_C}{\frac{\lambda^2}{k} + R_L^2}}$$

Hence, the power of this test is given by

$$(1-\beta) = Pr\left(T_2 \geq t_{1-\alpha,v'} | \phi, R_L, CV\right)$$

Unequal Variances Delta Method Z-Test

This procedure uses the following ratio hypotheses directly

$$H_0: \mu_T / \mu_C \leq R_L \quad \text{versus} \quad H_1: \mu_T / \mu_C > R_L$$

The null hypothesis about the ratio is tested using the delta method to determine the distribution of the ratio of two normal means. The unrestricted version of this test statistic is

$$T_4 = \frac{\frac{\bar{X}_T}{\bar{X}_C} - R_L}{\sqrt{\left(\frac{\bar{X}_T}{\bar{X}_C}\right)^2 \left(\frac{s_T^2}{\bar{X}_T^2 N_T} + \frac{s_C^2}{\bar{X}_C^2 N_C}\right)}}$$

The test assumes that T_4 is distributed as a standard normal distribution. Rothmann et al. (2012) state that the accuracy of the standard normal assumption depends on whether T_2 is standard normal and B is close to one, where

$$B = \frac{\sqrt{\frac{s_T^2}{N_T} + R_L^2 \frac{s_C^2}{N_C}}}{\sqrt{\frac{s_T^2}{N_T} + \left(\frac{\bar{X}_T}{\bar{X}_C}\right)^2 \frac{s_C^2}{N_C}}}$$

The power of this test is given by

$$z_\beta = \left(\frac{\phi - R_L}{CV}\right) \sqrt{\frac{N_C}{\frac{\lambda^2}{k} + \phi^2}} - z_\alpha$$

where T_4 is now assumed to follow the standard normal distribution mentioned above.

Example 1 – Finding Sample Size

A company has developed a generic drug for treating rheumatism and wants to show that it is not inferior to the standard drug. In this case, higher responses are better.

Responses are thought to follow a normal distribution with unequal variances. A parallel-group design will be used, and the data will be analyzed with a Satterthwaite corrected, two-sample t-test.

Researchers have decided to set the non-inferiority limit to 0.80. Past experience leads the researchers to set the CV to 1. The significance level is 0.025 and the power is 0.9. The sample size will be computed assuming that the mean ratio is 0.90, 0.95, or 1.00. The ratio of the two standard deviations is assumed to be 0.6, 0.8, or 1.0.

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**
Higher Means Are	**Better (H1: R > R_L, where R_L < 1)**
Power	**0.9**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Test Statistic	**Unequal Variances Satterthwaite T-Test**
R_L (Lower Non-Inferiority Limit)	**0.8**
R1 (Actual Mean Ratio, μ1 / μ2)	**0.9 0.95 1.0**
CV (Coef of Variation, σ2 / μ2)	**1**
λ (σ Ratio, σ1 / σ2)	**0.6 0.8 1**

Output

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

Numeric Reports

Numeric Results

Solve For:	Sample Size
Groups:	1 = Treatment, 2 = Control
Ratio:	R = μ1 / μ2
Higher Means Are:	Better
Hypotheses:	H0: R ≤ RL vs. H1: R > RL
Test:	Unequal Variances Satterthwaite T-Test

Power		Sample Size			Mean Ratio				
					Lower Non-Inferiority Limit	Actual	Control Group Coefficient of Variation	Standard Deviation Ratio	
Target	Actual	N1	N2	N	RL	R1	CV	λ	Alpha
0.9	0.90000	1051	1051	2102	0.8	0.90	1	0.6	0.025
0.9	0.90017	1346	1346	2692	0.8	0.90	1	0.8	0.025
0.9	0.90009	1724	1724	3448	0.8	0.90	1	1.0	0.025
0.9	0.90046	468	468	936	0.8	0.95	1	0.6	0.025
0.9	0.90001	598	598	1196	0.8	0.95	1	0.8	0.025
0.9	0.90033	767	767	1534	0.8	0.95	1	1.0	0.025
0.9	0.90030	264	264	528	0.8	1.00	1	0.6	0.025
0.9	0.90045	337	337	674	0.8	1.00	1	0.8	0.025
0.9	0.90063	432	432	864	0.8	1.00	1	1.0	0.025

Target Power	The desired power value (or values) entered in the procedure. Power is the probability of rejecting a false null hypothesis.
Actual Power	The power obtained in this scenario. Because N1 and N2 are discrete, this value is often (slightly) larger than the target power.
N1	The number of subjects sampled from the treatment population.
N2	The number of subjects sampled from the control population.
N	The total sample size. N = N1 + N2.
RL	The non-inferiority limit (or boundary) of the ratio. Since higher means are better, this value is less than one. This is the minimum that the ratio can be and still conclude that the treatment group is not inferior to the control group.
R1	The mean ratio (treatment/control) at which the power is computed.
CV	The coefficient of variation of the control group. CV = σ2 / μ2.
λ	The ratio of the standard deviations of the treatment and control groups. λ = σ1 / σ2.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design (where higher means are considered to be better) will be used to test whether the treatment mean (μ1) is non-inferior to the control (reference) mean (μ2), by testing whether the ratio of means (μ1 / μ2) is greater than the non-inferiority bound of 0.8 (H0: μ1 / μ2 ≤ 0.8 versus H1: μ1 / μ2 > 0.8). The comparison will be made using the original (untransformed) data with a two-sample, one-sided, unequal variances Satterthwaite t-test, with a Type I error rate (α) of 0.025. The ratio of the group standard deviations (σ1 / σ2) is assumed to be 0.6, and the coefficient of variation of the control group (σ2 / μ2) is assumed to be 1. To detect a mean ratio of 0.9 with 90% power, the number of subjects needed will be 1051 in Group 1 (treatment), and 1051 in Group 2 (control) (a total of 2102 subjects).

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%	1051	1051	2102	1314	1314	2628	263	263	526
20%	1346	1346	2692	1683	1683	3366	337	337	674
20%	1724	1724	3448	2155	2155	4310	431	431	862
20%	468	468	936	585	585	1170	117	117	234
20%	598	598	1196	748	748	1496	150	150	300
20%	767	767	1534	959	959	1918	192	192	384
20%	264	264	528	330	330	660	66	66	132
20%	337	337	674	422	422	844	85	85	170
20%	432	432	864	540	540	1080	108	108	216

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, 1314 subjects should be enrolled in Group 1, and 1314 in Group 2, to obtain final group sample sizes of 1051 and 1051, respectively.

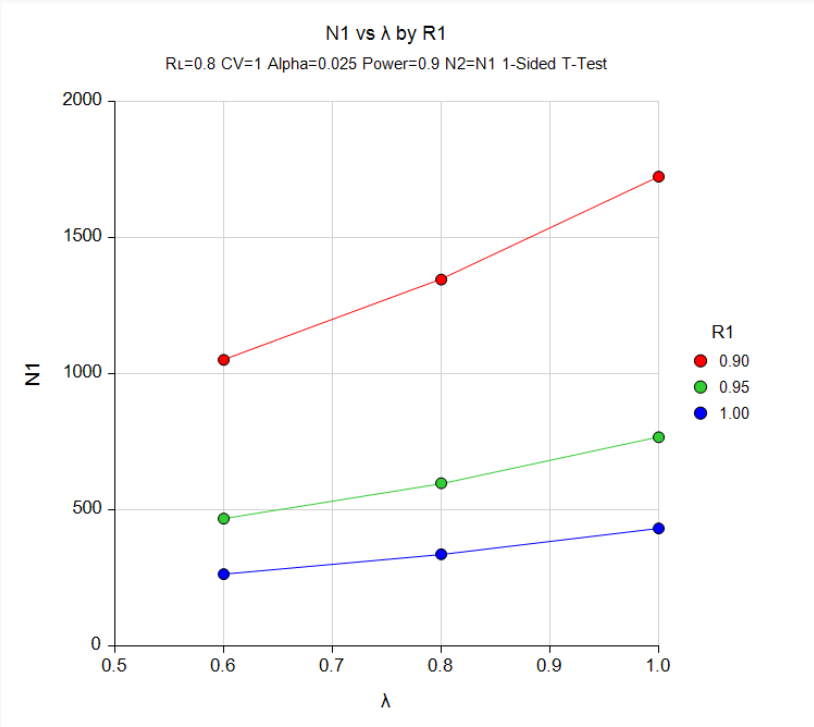
References

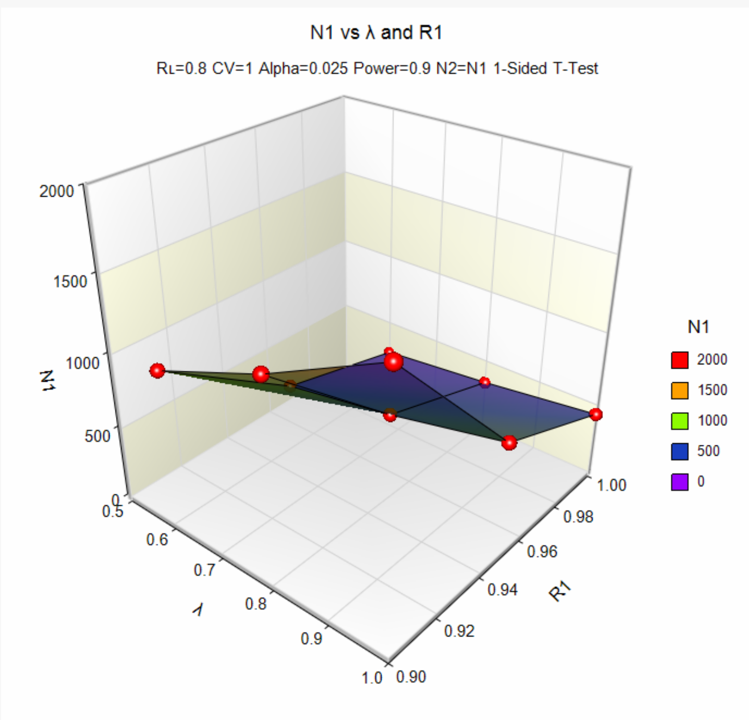
Rothmann, M.D., Wiens, B.L., and Chan, I.S.F. 2012. Design and Analysis of Non-Inferiority Trials. Taylor & Francis/CRC Press. Boca Raton, Florida.
Kieser, M. and Hauschke, D. 1999. 'Approximate Sample Sizes for Testing Hypotheses about the Ratio and Difference of Two Means.' Journal of Biopharmaceutical 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.

This report shows the sample size required for the indicated scenarios.

Plots Section

Plots





These plots show the sample size for the various scenarios.

Example 2 – Validation using Rothmann (2012)

Rothmann *et al.* (2012) present a table on page 342 in which they calculate several sample sizes. Specifically, they calculate the sample size for the large sample z-test to be 20 in each group.

The non-inferiority limit is 0.75, the CV is 0.3, the significance level is 0.025, the power is 0.9, the SD ratio is 0.5, and mean ratio is 0.95.

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**
Higher Means Are	**Better (H1: R > RL, where RL < 1)**
Power	**0.9**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Test Statistic	**Unequal Variances Large Sample Z-Test**
RL (Lower Non-Inferiority Limit)	**0.75**
R1 (Actual Mean Ratio, μ1 / μ2)	**0.95**
CV (Coef of Variation, σ2 / μ2)	**0.3**
λ (σ Ratio, σ1 / σ2)	**0.5**

Output

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

Numeric Results

Solve For:	Sample Size
Groups:	1 = Treatment, 2 = Control
Ratio:	R = μ1 / μ2
Higher Means Are:	Better
Hypotheses:	H0: R ≤ RL vs. H1: R > RL
Test:	Unequal Variances Large Sample Z-Test

Power		Sample Size			Mean Ratio		Control Group Coefficient of Variation	Standard Deviation Ratio	
Target	Actual	N1	N2	N	Lower Non-Inferiority Limit RL	Actual R1	CV	λ	Alpha
0.9	0.91111	20	20	40	0.75	0.95	0.3	0.5	0.025

PASS also calculates the sample size in each group to be 20.