

Chapter 573

Superiority by a Margin Tests for the Ratio of Two Means (Normal Data)

Introduction

This procedure calculates power and sample size for *superiority by a margin* 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 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 *Superiority by a Margin Tests for the Ratio of Two Means (Log-Normal Data)* to obtain more accurate results.

Superiority 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	μ_1	<i>Treatment mean.</i> This is the treatment mean.
μ_C	μ_2	<i>Control (Reference) mean.</i> This is the mean of a reference population.
R_L	R_L	<i>Lower Superiority Limit.</i> This is the upper limit for the mean ratio when higher values are 'worse'. Values below this amount are assumed to be superior. Values above this amount as assumed to be non-superior.
R_U	R_U	<i>Upper Superiority Limit.</i> This is the upper limit for the mean ratio when lower values are 'worse. Values above this amount are assumed to be superior. Values below this amount as assumed to be non-superior.
ϕ	R1	<i>Actual ratio.</i> 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.

Superiority by a Margin Tests for the Ratio of Two Means (Normal Data)

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 superiority limit. The value of ϕ at which power is calculated should be greater than R_U .

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

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 the superiority limit. The value of ϕ at which power is calculated must be less than R_L .

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

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 equivalence 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 superiority by a margin 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. Also, assume that larger responses are 'better'.

Assuming that $R_U > 1$, the superiority by a margin hypotheses are

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

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

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

to

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

The null hypothesis is tested using the test statistic

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

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

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

Hence, the power of this test is given by

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

Unequal Variances Large Sample Z-Test

The ratio hypotheses are rearranged as from

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

to

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

The null hypothesis is tested using the test statistic

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

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 superiority by a margin is concluded at a one-sided level of α .

Hence, the approximate power of this test is given by

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

This can be rearranged to give

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

Unequal Variances Satterthwaite T-Test

The ratio hypotheses are rearranged as from

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

to

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

The null hypothesis is tested using the test statistic

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

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 superiority by a margin is concluded at a one-sided level of α . The Satterthwaite degrees of freedom is given by

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

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

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

The non-centrality parameter is given by

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

Hence, the power of this test is given by

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

Unequal Variances Delta Method Z-Test

This procedure uses the following ratio hypotheses directly

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

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

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

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

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

The power of this test is given by

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

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 superior 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 superiority limit to 1.25. 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 1.5, 1.6, or 1.7. 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: $R > R_u$, where $R_u > 1$)
Power	0.90
Alpha	0.025
Group Allocation	Equal ($N_1 = N_2$)
Test Statistic	Unequal Variances Satterthwaite T-Test
R_u (Superiority Limit)	1.25
R_1 (Actual Mean Ratio, μ_1 / μ_2)	1.5 1.6 1.7
CV (Coef of Variation, σ_2 / μ_2)	1
λ (σ Ratio, σ_1 / σ_2)	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: $R = \mu_1 / \mu_2$
 Higher Means Are: Better
 Hypotheses: $H_0: R \leq R_u$ vs. $H_1: R > R_u$
 Test: Unequal Variances Satterthwaite T-Test

Power		Sample Size			Mean Ratio		Control Group Coefficient of Variation CV	Standard Deviation Ratio λ	Alpha
					Upper Superiority Limit R_u	Actual R_1			
Target	Actual	N1	N2	N					
0.9	0.90040	325	325	650	1.25	1.5	1	0.6	0.025
0.9	0.90034	371	371	742	1.25	1.5	1	0.8	0.025
0.9	0.90063	432	432	864	1.25	1.5	1	1.0	0.025
0.9	0.90130	167	167	334	1.25	1.6	1	0.6	0.025
0.9	0.90142	191	191	382	1.25	1.6	1	0.8	0.025
0.9	0.90025	221	221	442	1.25	1.6	1	1.0	0.025
0.9	0.90254	102	102	204	1.25	1.7	1	0.6	0.025
0.9	0.90143	116	116	232	1.25	1.7	1	0.8	0.025
0.9	0.90005	134	134	268	1.25	1.7	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. $N = N_1 + N_2$.

R_u The superiority limit (or boundary) of the ratio. Since higher means are better, this value is greater than one. This is the smallest that the ratio can be and still conclude that the treatment group is superior to the control group.

R_1 The mean ratio (treatment/control) at which the power is computed.

CV The coefficient of variation of the control group. $CV = \sigma_2 / \mu_2$.

λ The ratio of the standard deviations of the treatment and control groups. $\lambda = \sigma_1 / \sigma_2$.

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 (μ_1) is superior to the control (reference) mean (μ_2) by a margin, by testing whether the ratio of means (μ_1 / μ_2) is greater than the superiority bound of 1.25 ($H_0: \mu_1 / \mu_2 \leq 1.25$ versus $H_1: \mu_1 / \mu_2 > 1.25$). 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 (σ_1 / σ_2) is assumed to be 0.6, and the coefficient of variation of the control group (σ_2 / μ_2) is assumed to be 1. To detect a mean ratio of 1.5 with 90% power, the number of subjects needed will be 325 in Group 1 (treatment), and 325 in Group 2 (control) (a total of 650 subjects).

Superiority by a Margin Tests for the Ratio of Two Means (Normal Data)

Dropout-Inflated Sample Size

Dropout Rate	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
	N1	N2	N	N1'	N2'	N'	D1	D2	D
20%	325	325	650	407	407	814	82	82	164
20%	371	371	742	464	464	928	93	93	186
20%	432	432	864	540	540	1080	108	108	216
20%	167	167	334	209	209	418	42	42	84
20%	191	191	382	239	239	478	48	48	96
20%	221	221	442	277	277	554	56	56	112
20%	102	102	204	128	128	256	26	26	52
20%	116	116	232	145	145	290	29	29	58
20%	134	134	268	168	168	336	34	34	68

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, 407 subjects should be enrolled in Group 1, and 407 in Group 2, to obtain final group sample sizes of 325 and 325, respectively.

References

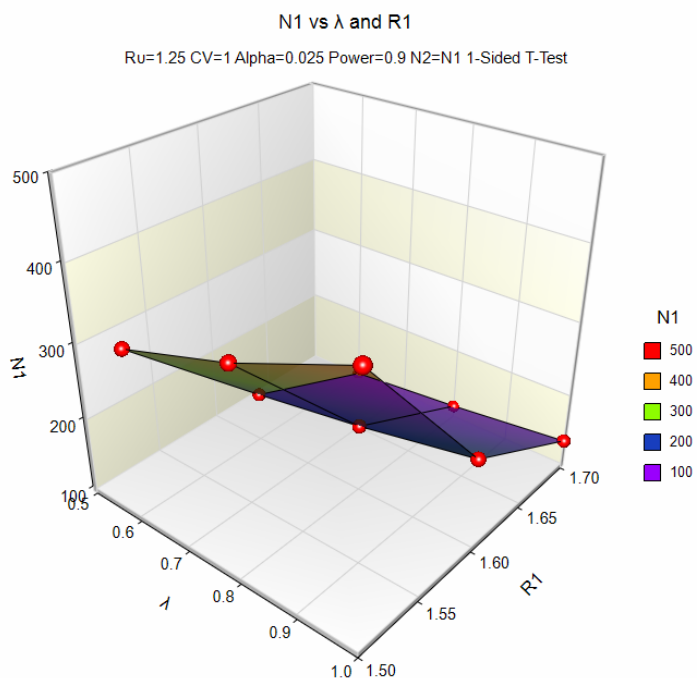
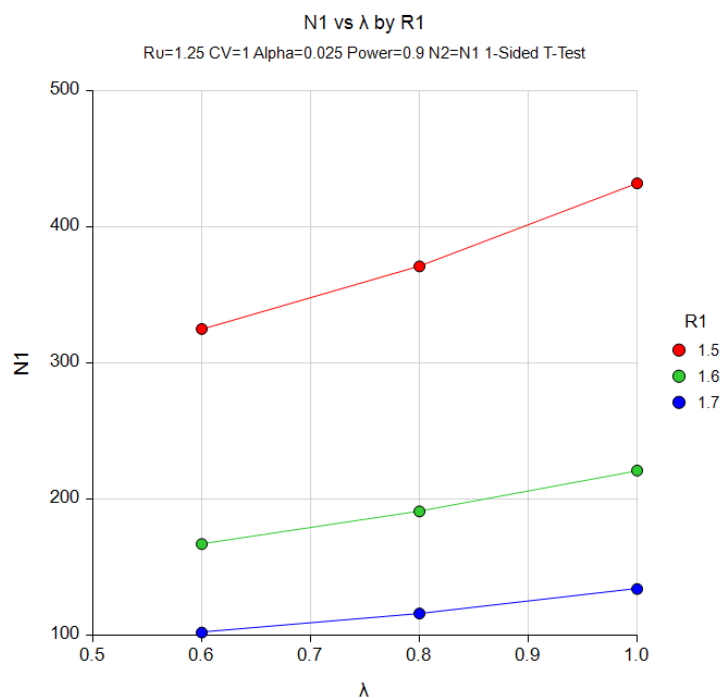
- Rothmann, M.D., Wiens, B.L., and Chan, I.S.F. 2012. Design and Analysis of Non-Inferiority Trials. Taylor & Francis/CRC Press. Boca Raton, Florida.
- Kieser, M. and Hauschke, D. 1999. 'Approximate Sample Sizes for Testing Hypotheses about the Ratio and Difference of Two Means.' Journal of Biopharmaceutical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

Superiority by a Margin Tests for the Ratio of Two Means (Normal Data)

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. Note that these results were for non-inferiority testing, but the formulas work for superiority tests as well.

The superiority 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 the 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: $R > R_u$, where $R_u > 1$)
Power.....	0.90
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Test Statistic	Unequal Variances Large Sample Z-Test
R_u (Superiority Limit)	0.75
R_1 (Actual Mean Ratio, μ_1 / μ_2)	0.95
CV (Coef of Variation, σ_2 / μ_2).....	0.3
λ (σ Ratio, σ_1 / σ_2).....	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:	$R = \mu_1 / \mu_2$
Higher Means Are:	Better
Hypotheses:	$H_0: R \leq R_u$ vs. $H_1: R > R_u$
Test:	Unequal Variances Large Sample Z-Test

Power		Sample Size			Mean Ratio		Control Group Coefficient of Variation CV	Standard Deviation Ratio λ	Alpha
		N1	N2	N	Upper Superiority Limit R_u	Actual R_1			
Target	Actual								
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.