Chapter 455

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

Introduction

This procedure calculates power and sample size for *non-inferiority* t-tests from a parallel-groups design in which the logarithm of the outcome is a continuous normal random variable. 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 testing the non-inferiority of two treatments using data from a two-group design are given in another chapter and they will not be repeated here. If the logarithm of the response can be assumed to follow a normal distribution, hypotheses about non-inferiority stated in terms of the ratio can be transformed into hypotheses about the difference. The details of this analysis are given in Julious (2004). They will only be summarized here.

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_R	Not used	*Reference mean*. This is the mean of a reference population.
M_{NI}	NIM	*Margin of non-inferiority.* This is a tolerance value that defines the maximum amount that is not of practical importance. This is the largest change in the mean ratio from the baseline value (usually one) that is still considered to be trivial.
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 one minus the margin of non-inferiority. The value of ϕ at which power is calculated must be greater than $\phi_0 = 1 - |M_{NI}|$.

$H_0: \phi \leq 1 - |M_{NI}|$ versus $H_1: \phi > 1 - |M_{NI}|$

$H_0: \phi \leq \phi_0$ versus $H_1: \phi > \phi_0$

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 $\phi_0 = 1 + |M_{NI}|$.

$H_0: \phi \geq 1 + |M_{NI}|$ versus $H_1: \phi < 1 + |M_{NI}|$

$H_0: \phi \geq \phi_0$ versus $H_1: \phi < \phi_0$

Log-Transformation

In many cases, hypotheses stated in terms of ratios are more convenient than hypotheses stated in terms of differences. This is because ratios can be interpreted as scale-less percentages, but differences must be interpreted as actual amounts in their original scale. Hence, it has become a common practice to take the following steps in hypothesis testing.

1. State the statistical hypotheses in terms of ratios.
2. Transform these into hypotheses about differences by taking logarithms.
3. Analyze the logged data—that is, do the analysis in terms of the difference.
4. Draw the conclusion in terms of the ratio.

The details of step 2 for the null hypothesis when higher means are better are as follows:

$$H_0: \phi \leq \phi_0 \quad \Rightarrow \quad H_0: \frac{\mu_T}{\mu_R} \leq \phi_0 \quad \Rightarrow \quad H_0: \ln(\mu_T) - \ln(\mu_R) \leq \ln(\phi_0)$$

Thus, a hypothesis about the ratio of the means on the original scale can be translated into a hypothesis about the difference of two means on the logged scale.

Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean. This parameter can be used to represent the variation in the data because of a unique relationship that it has in the case of log-normal data.

Suppose the variable X is the logarithm of the original variable Y. That is, $X = \ln(Y)$ and $Y = \exp(X)$. Label the mean and variance of X as μ_X and σ_X^2, respectively. Similarly, label the mean and variance of Y as μ_Y and σ_Y^2, respectively. If X is normally distributed, then Y is log-normally distributed. Julious (2004) presents the following well-known relationships between these two variables

$$\mu_Y = e^{\mu_X + \frac{\sigma_X^2}{2}}$$

$$\sigma_Y^2 = \mu_Y^2 \left(e^{\sigma_X^2} - 1 \right)$$

From this relationship, the coefficient of variation of *Y* can be found to be

$$COV_Y = \frac{\sqrt{\mu_Y^2\left(e^{\sigma_X^2} - 1\right)}}{\mu_Y}$$

$$= \sqrt{e^{\sigma_X^2} - 1}$$

Solving this relationship for σ_X^2, the standard deviation of *X* can be stated in terms of the coefficient of variation of *Y* as

$$\sigma_X = \sqrt{\ln\left(COV_Y^2 + 1\right)}$$

Similarly, the mean of *X* is

$$\mu_X = \ln\left(\frac{\mu_Y}{\sqrt{COV_Y^2 + 1}}\right)$$

Thus, the hypotheses can be stated in the original (Y) scale and then the power can be analyzed in the transformed (X) scale. For parallel-group designs, $\sigma_X^2 = \sigma_d^2$, the average variance used in the t-test of the logged data.

Power Calculation

As is shown above, the hypotheses can be stated in the original (Y) scale using ratios or the logged (X) scale using differences. In either case, the power and sample size calculations are made using the formulas for testing the difference in two means. These formulas are presented in another chapter and are not duplicated here.

Example 1 – Finding Power

A company has developed a generic drug for treating rheumatism and wants to show that it is not inferior to the standard drug. Responses following either treatment are known to follow a log normal distribution. A parallel-group design will be used and the logged data will be analyzed with a two-sample t-test.

Researchers have decided to set the margin of equivalence at 0.20. Past experience leads the researchers to set the COV to 1.50. The significance level is 0.025. The power will be computed assuming that the true ratio is either 0.95 or 1.00. Sample sizes between 100 and 1000 will be included in the analysis.

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	**Power**
Higher Means Are...	**Better (H1: R > 1 - NIM)**
Alpha...	**0.025**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**100 to 1000 by 100**
NIM (Non-Inferiority Margin)	**0.2**
R1 (Actual Ratio)	**0.95 1.0**
COV (Coefficient of Variation).......................	**1.5**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Reference
Ratio: R = Treatment Mean / Reference Mean
Higher Means Are: Better
Hypotheses: H0: R ≤ 1 - NIM vs. H1: R > 1 - NIM

	Sample Size			Non-Inferiority		Actual	Coefficient	
Power	N1	N2	N	Margin -NIM	Bound R0	Ratio R1	of Variation COV	Alpha
0.19875	100	100	200	-0.2	0.8	0.95	1.5	0.025
0.35165	200	200	400	-0.2	0.8	0.95	1.5	0.025
0.49026	300	300	600	-0.2	0.8	0.95	1.5	0.025
0.60984	400	400	800	-0.2	0.8	0.95	1.5	0.025
0.70637	500	500	1000	-0.2	0.8	0.95	1.5	0.025
0.78275	600	600	1200	-0.2	0.8	0.95	1.5	0.025
0.84160	700	700	1400	-0.2	0.8	0.95	1.5	0.025
0.88598	800	800	1600	-0.2	0.8	0.95	1.5	0.025
0.91886	900	900	1800	-0.2	0.8	0.95	1.5	0.025
0.94283	1000	1000	2000	-0.2	0.8	0.95	1.5	0.025
0.30375	100	100	200	-0.2	0.8	1.00	1.5	0.025
0.53604	200	200	400	-0.2	0.8	1.00	1.5	0.025
0.70997	300	300	600	-0.2	0.8	1.00	1.5	0.025
0.82799	400	400	800	-0.2	0.8	1.00	1.5	0.025
0.90132	500	500	1000	-0.2	0.8	1.00	1.5	0.025
0.94511	600	600	1200	-0.2	0.8	1.00	1.5	0.025
0.97024	700	700	1400	-0.2	0.8	1.00	1.5	0.025
0.98421	800	800	1600	-0.2	0.8	1.00	1.5	0.025
0.99178	900	900	1800	-0.2	0.8	1.00	1.5	0.025
0.99579	1000	1000	2000	-0.2	0.8	1.00	1.5	0.025

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1 and N2 The number of items sampled from each population.
N The total sample size. N = N1 + N2.
-NIM The magnitude and direction of the margin of non-inferiority. Since higher means are better, this value is negative and is the distance below one that the ratio can be and still conclude non-inferiority.
R0 The non-inferiority bound on the ratio corresponding to NIM. R0 = 1 - NIM.
R1 The mean ratio (treatment/reference) at which the power is computed.
COV The coefficient of variation on the original scale.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design will be used to test whether the Group 1 (treatment) mean is non-inferior to the Group 2 (reference) mean, with a non-inferiority margin (difference from a ratio of 1) of -0.2 (H0: μ1 / μ2 ≤ 0.8 versus Ha: μ1 / μ2 > 0.8). The comparison will be made using a one-sided, two-sample t-test using a log-transformation, with a Type I error rate (α) of 0.025. The coefficient of variation on the original scale for both groups is assumed to be 1.5. To detect a ratio of means (μ1 / μ2) of 0.95, with a sample size of 100 in Group 1 and 100 in Group 2, the power is 0.19875.

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%	100	100	200	125	125	250	25	25	50
20%	200	200	400	250	250	500	50	50	100
20%	300	300	600	375	375	750	75	75	150
20%	400	400	800	500	500	1000	100	100	200
20%	500	500	1000	625	625	1250	125	125	250
20%	600	600	1200	750	750	1500	150	150	300
20%	700	700	1400	875	875	1750	175	175	350
20%	800	800	1600	1000	1000	2000	200	200	400
20%	900	900	1800	1125	1125	2250	225	225	450
20%	1000	1000	2000	1250	1250	2500	250	250	500

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 (as entered by the user). 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. 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, 125 subjects should be enrolled in Group 1, and 125 in Group 2, to obtain final group sample sizes of 100 and 100, respectively.

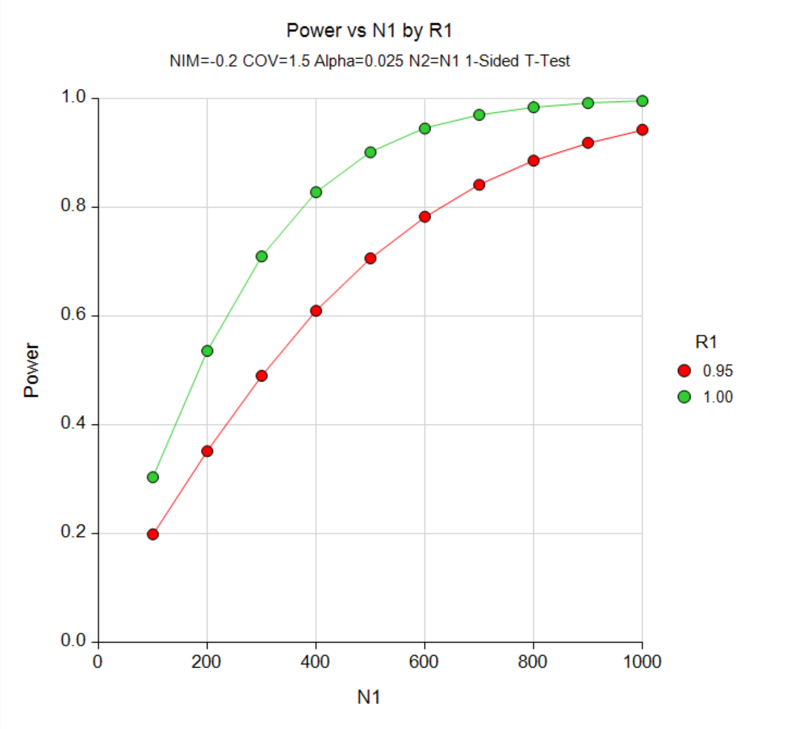
References

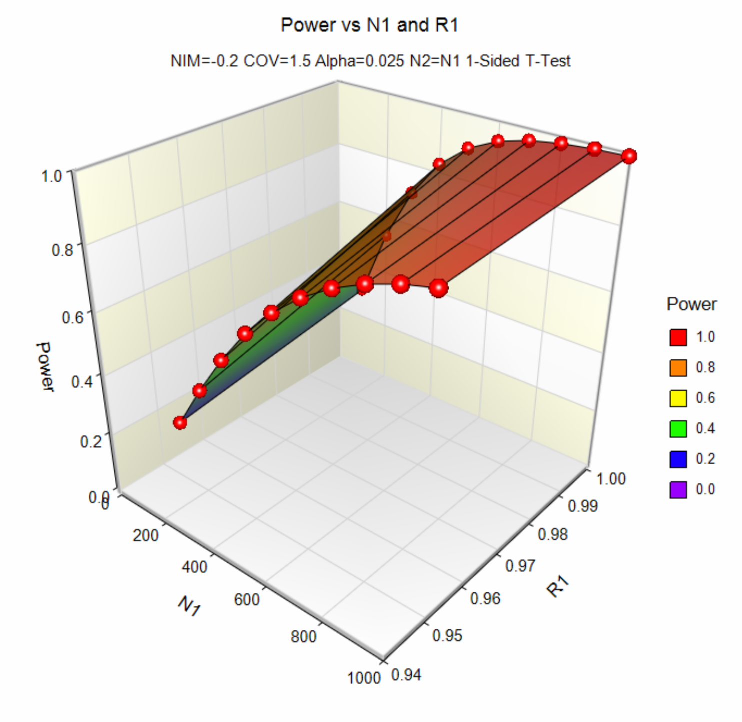
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.
Julious, Steven A. 2004. 'Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data.' Statistics in Medicine, 23:1921-1986.

This report shows the power for the indicated scenarios.

Plots Section

Plots





These plots show the power versus the sample size for two R1 values.

Example 2 – Validation

We could not find a validation example for this procedure in the statistical literature. Therefore, we will show that this procedure gives the same results as the non-inferiority test on differences—a procedure that has been validated (Two-Sample T-Tests for Non-Inferiority Assuming Equal Variance). We will use the same settings as those given in Example1. Since the output for this example is shown above, all that we need is the output from the procedure that uses differences.

To run the inferiority test on differences, we need the values of NIM, δ, and σ.

$$NIM' = \ln(1 - NIM)$$

$$= \ln(0.8)$$

$$= 0.223144$$

$$\delta = \ln(R1)$$

$$= \ln(0.95)$$

$$= -0.051293$$

$$\sigma = \sqrt{\ln(COV_Y^2 + 1)}$$

$$= \ln(1.5^2 + 1)$$

$$= 1.085659$$

Setup

If the procedure window (Two-Sample T-Tests for Non-Inferiority Assuming Equal Variance) 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**
Higher Means Are	**Better (H1: δ > -NIM)**
Alpha	**0.025**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**100 to 1000 by 100**
NIM (Non-Inferiority Margin)	**0.223144**
δ (Actual Difference to Detect)	**-0.051293 0.0**
σ (Standard Deviation)	**1.085659**

Output

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

Numeric Results

Solve For: Power
Test Type: Two-Sample Equal-Variance T-Test
Difference: δ = μ1 - μ2 = μT - μR
Higher Means Are: Better
Hypotheses: H0: δ ≤ -NIM vs. H1: δ > -NIM

Power	Sample Size N1	N2	N	Non-Inferiority Margin -NIM	Mean Difference δ	Standard Deviation σ	Alpha
0.19875	100	100	200	-0.223144	-0.051293	1.085659	0.025
0.35165	200	200	400	-0.223144	-0.051293	1.085659	0.025
0.49026	300	300	600	-0.223144	-0.051293	1.085659	0.025
0.60984	400	400	800	-0.223144	-0.051293	1.085659	0.025
0.70637	500	500	1000	-0.223144	-0.051293	1.085659	0.025
0.78275	600	600	1200	-0.223144	-0.051293	1.085659	0.025
0.84160	700	700	1400	-0.223144	-0.051293	1.085659	0.025
0.88599	800	800	1600	-0.223144	-0.051293	1.085659	0.025
0.91886	900	900	1800	-0.223144	-0.051293	1.085659	0.025
0.94284	1000	1000	2000	-0.223144	-0.051293	1.085659	0.025
0.30375	100	100	200	-0.223144	0.000000	1.085659	0.025
0.53604	200	200	400	-0.223144	0.000000	1.085659	0.025
0.70997	300	300	600	-0.223144	0.000000	1.085659	0.025
0.82799	400	400	800	-0.223144	0.000000	1.085659	0.025
0.90132	500	500	1000	-0.223144	0.000000	1.085659	0.025
0.94511	600	600	1200	-0.223144	0.000000	1.085659	0.025
0.97024	700	700	1400	-0.223144	0.000000	1.085659	0.025
0.98421	800	800	1600	-0.223144	0.000000	1.085659	0.025
0.99178	900	900	1800	-0.223144	0.000000	1.085659	0.025
0.99579	1000	1000	2000	-0.223144	0.000000	1.085659	0.025

You can compare these power values with those shown above in Example 1 to validate the procedure. You will find that the power values are identical.