

Chapter 455

Non-Inferiority Tests for the Ratio of Two Means

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.

<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
μ_T	Not used	<i>Treatment mean.</i> This is the treatment mean.
μ_R	Not used	<i>Reference mean.</i> This is the mean of a reference population.
M_{NI}	NIM	<i>Margin of non-inferiority.</i> 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.
<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
ϕ	R1	<i>True 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.

The null hypothesis of inferiority is

$$H_0: \phi \leq \phi_L \quad \text{where } \phi_L < 1.$$

Non-Inferiority Tests for the Ratio of Two Means

and the alternative hypothesis of non-inferiority is

$$H_1: \phi > \phi_L$$

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 are as follows.

$$\begin{aligned} \phi_L &\leq \phi \\ \Rightarrow \phi_L &\leq \left\{ \frac{\mu_T}{\mu_R} \right\} \\ \Rightarrow \ln(\phi_L) &\leq \{ \ln(\mu_T) - \ln(\mu_R) \} \end{aligned}$$

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

$$\begin{aligned} \mu_Y &= \left(e^{\mu_X + \frac{\sigma_X^2}{2}} \right) \\ \sigma_Y^2 &= \mu_Y^2 \left(e^{\sigma_X^2} - 1 \right) \end{aligned}$$

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

$$\begin{aligned} COV_Y &= \frac{\sqrt{\mu_Y^2 (e^{\sigma_X^2} - 1)}}{\mu_Y} \\ &= \sqrt{e^{\sigma_X^2} - 1} \end{aligned}$$

Non-Inferiority Tests for the Ratio of Two Means

Solving this relationship for σ_X^2 , the standard deviation of X can be stated in terms of the coefficient of variation of Y . This equation is

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

Similarly, the mean of X is

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

One final note: for parallel-group designs, σ_X^2 equals σ_d^2 , the average variance used in the t-test of the logged data.

Thus, the hypotheses can be stated in the original (Y) scale and then the power can be analyzed in the transformed (X) scale.

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.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains the parameters associated with this test such as the means, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. In most situations, you will select either *Power* or *Sample Size (NI)*.

Test

Higher Means Are

This option defines whether higher values of the response variable are to be considered better or worse. The choice here determines the direction of the non-inferiority test.

If Higher Means Are Better the null hypothesis is $R \leq 1$ -NIM and the alternative hypothesis is $R > 1$ -NIM. If Higher Means Are Worse the null hypothesis is $R \geq 1$ +NIM and the alternative hypothesis is $R < 1$ +NIM.

Non-Inferiority Tests for the Ratio of Two Means

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of inferiority when in fact the treatment mean is non-inferior.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. In this procedure, a type-I error occurs when rejecting the null hypothesis of inferiority when in fact the treatment group is not inferior to the reference group.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on $N1$ or $N2$ or both.

The options are

- **Equal ($N1 = N2$)**
This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.
- **Enter $N2$, solve for $N1$**
Select this option when you wish to fix $N2$ at some value (or values), and then solve only for $N1$. Please note that for some values of $N2$, there may not be a value of $N1$ that is large enough to obtain the desired power.
- **Enter $R = N2/N1$, solve for $N1$ and $N2$**
For this choice, you set a value for the ratio of $N2$ to $N1$, and then PASS determines the needed $N1$ and $N2$, with this ratio, to obtain the desired power. An equivalent representation of the ratio, R , is

$$N2 = R * N1.$$
- **Enter percentage in Group 1, solve for $N1$ and $N2$**
For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed $N1$ and $N2$ with this percentage to obtain the desired power.

$N2$ (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter $N2$, solve for $N1$ "

$N2$ is the number of items or individuals sampled from the Group 2 population.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

Non-Inferiority Tests for the Ratio of Two Means

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N2/N1, solve for N1 and N2."

R is the ratio of N2 to N1. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of N2 to N1 while solving for N1 and N2. Only sample size combinations with this ratio are considered.

N2 is related to N1 by the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N1 = 10$ and $N2 = 20$, or $N1 = 50$ and $N2 = 100$).

R must be greater than 0. If $R < 1$, then N2 will be less than N1; if $R > 1$, then N2 will be greater than N1. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N1 and N2."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for N1 and N2. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When Not Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**
This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.
- **Enter N1 and N2 individually**
This choice permits you to enter different values for N1 and N2.
- **Enter N1 and R, where N2 = R * N1**
Choose this option to specify a value (or values) for N1, and obtain N2 as a ratio (multiple) of N1.
- **Enter total sample size and percentage in Group 1**
Choose this option to specify a value (or values) for the total sample size (N), obtain N1 as a percentage of N, and then N2 as $N - N1$.

Non-Inferiority Tests for the Ratio of Two Means

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal ($N1 = N2$)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for $N1$, and also the value for $N2$.

The Sample Size Per Group must be ≥ 2 . You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter $N1$ and $N2$ individually" or "Enter $N1$ and R , where $N2 = R * N1$."*

$N1$ is the number of items or individuals sampled from the Group 1 population.

$N1$ must be ≥ 2 . You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter $N1$ and $N2$ individually."

$N2$ is the number of items or individuals sampled from the Group 2 population.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter $N1$ and R , where $N2 = R * N1$."*

R is the ratio of $N2$ to $N1$. That is,

$$R = N2/N1$$

Use this value to obtain $N2$ as a multiple (or proportion) of $N1$.

$N2$ is calculated from $N1$ using the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines $N1$ and $N2$.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Non-Inferiority Tests for the Ratio of Two Means

Effect Size – Ratios

NIM (Non-Inferiority Margin)

This is the magnitude of the margin of non-inferiority. It must be entered as a positive number.

When higher means are better, this value is the distance below one for which the mean ratio (Treatment Mean / Reference Mean) still indicates non-inferiority of the treatment mean. E.g., a value of 0.2 here specifies that mean ratios greater than 0.8 indicate non-inferiority of the treatment mean.

When higher means are worse, this value is the distance above one for which the mean ratio (Treatment Mean / Reference Mean) still indicates non-inferiority of the treatment mean. E.g., a value of 0.2 here specifies that mean ratios less than 1.2 indicate non-inferiority of the treatment mean.

Note that the sign of this value is ignored. Only the magnitude is used.

Recommended values:

0.20 is a common value for this parameter.

R1 (True Ratio)

This is the value of the ratio of the two means (Treatment Mean / Reference Mean) at which the power is to be calculated.

Often, the ratio will be set to one. However, some authors recommend using a ratio slightly different than one, such as 0.95 (when higher means are "better") or 1.05 (when higher means are "worse"), since this will require a larger sample size.

Effect Size – Coefficient of Variation

COV (Coefficient of Variation)

The coefficient of variation is used to specify the variability (standard deviation). It is important to realize that this is the COV defined on the original (not logged) scale. This value must be determined from past experience or from a pilot study. See the discussion above for more details on the definition of the coefficient of variation.

Non-Inferiority Tests for the Ratio of Two Means

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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for the Ratio of Two Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Means**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Higher Means Are.....	Better
Alpha	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group.....	100 to 1000 by 100
E (Equivalence Margin).....	0.20
R1 (True Ratio).....	0.95 1.0
COV (Coefficient of Variation).....	1.50

Annotated Output

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

Numeric Results

Numeric Results for Non-Inferiority Test (H0: $R \leq 1$ - NIM; H1: $R > 1$ - NIM)

Higher Means are Better

Test Statistic: T-Test

H0 (null hypothesis) is that $R \leq 1$ - NIM, where $R = \text{Treatment Mean} / \text{Reference Mean}$.

H1 (alternative hypothesis) is that $R > 1$ - NIM.

Power	N1	N2	N	-NIM	Bound (NIB)	R1	COV	Alpha
0.1987	100	100	200	-0.20	0.800	1	1.5	0.025
0.3516	200	200	400	-0.20	0.800	1	1.5	0.025
0.4903	300	300	600	-0.20	0.800	1	1.5	0.025
0.6098	400	400	800	-0.20	0.800	1	1.5	0.025
0.7064	500	500	1000	-0.20	0.800	1	1.5	0.025
0.7827	600	600	1200	-0.20	0.800	1	1.5	0.025
0.8416	700	700	1400	-0.20	0.800	1	1.5	0.025
0.8860	800	800	1600	-0.20	0.800	1	1.5	0.025
0.9189	900	900	1800	-0.20	0.800	1	1.5	0.025
0.9428	1000	1000	2000	-0.20	0.800	1	1.5	0.025
0.3038	100	100	200	-0.20	0.800	1	1.5	0.025
0.5360	200	200	400	-0.20	0.800	1	1.5	0.025
0.7100	300	300	600	-0.20	0.800	1	1.5	0.025

(report continues)

Non-Inferiority Tests for the Ratio of Two Means

References

Chow, S.C.; Shao, J.; Wang, H. 2003. Sample Size Calculations in Clinical Research. Marcel Dekker. New York.
 Julious, Steven A. 2004. 'Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data.'
 Statistics in Medicine, 23:1921-1986.

Report Definitions

Power is the probability of rejecting a false null hypothesis.

N_1 and N_2 are the number of items sampled from each population.

N is the total sample size, $N_1 + N_2$.

-NIM is 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 is still considered non-inferior. NIB is the corresponding bound to the non-inferiority margin, and equals $1 - NIM$.

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

COV is the coefficient of variation on the original scale.

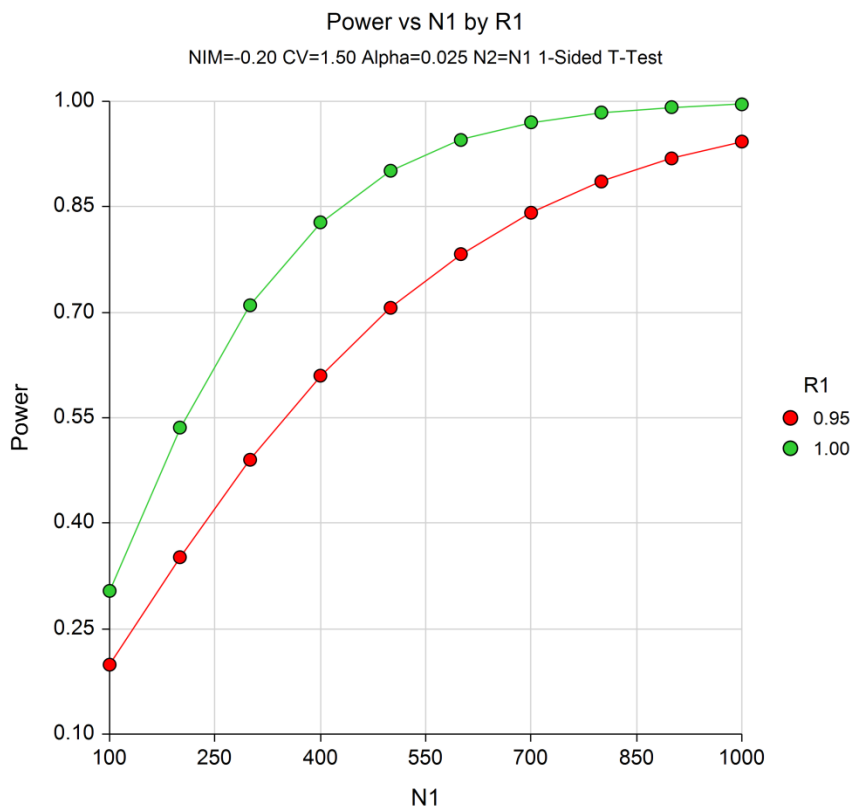
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

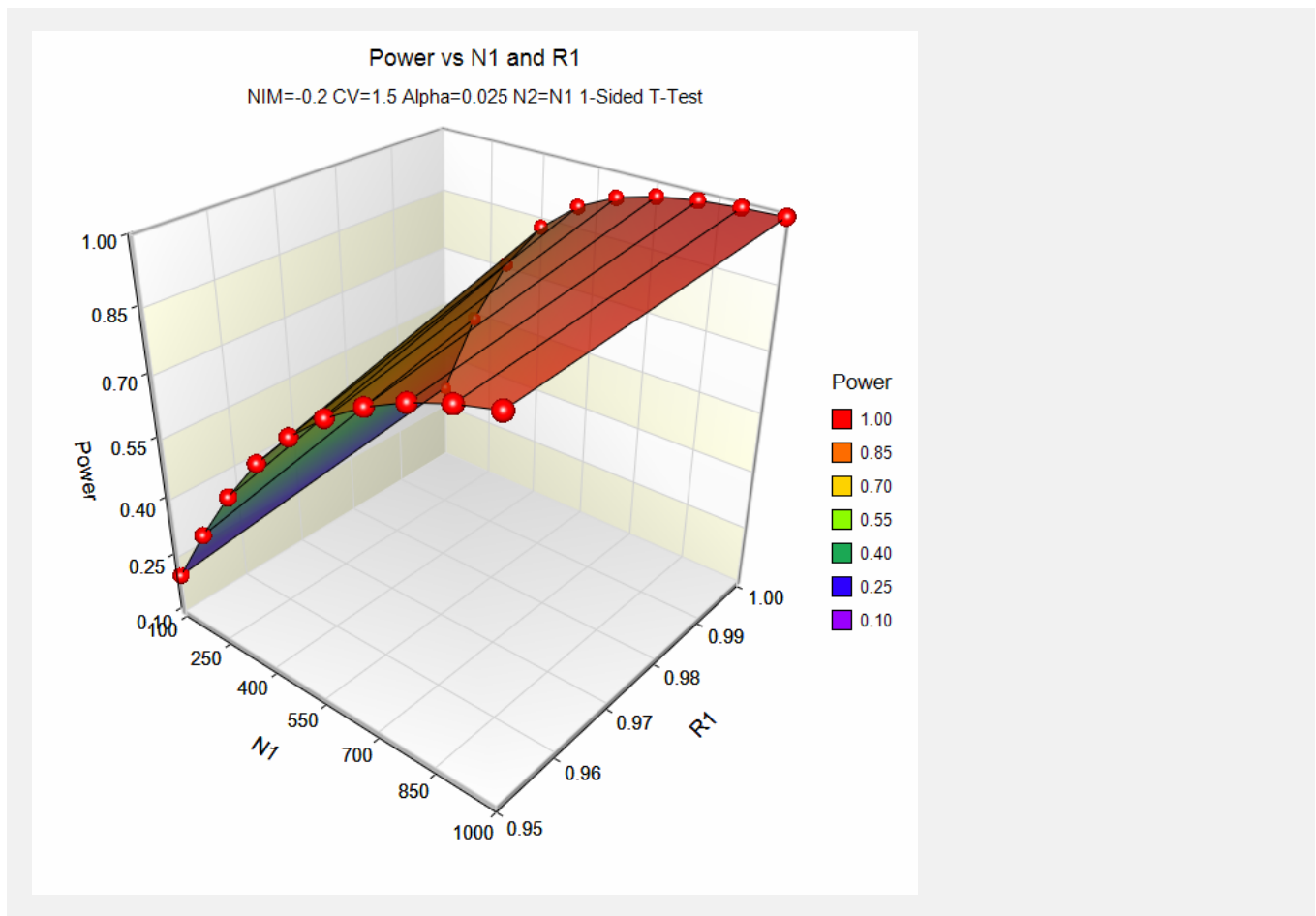
Group sample sizes of 100 in the first group and 100 in the second group achieve 20% power to detect non-inferiority using a one-sided, two-sample t-test. The margin of non-inferiority is -0.20. The true ratio of the means at which the power is evaluated is 1. The significance level (alpha) of the test is 0.025. The coefficients of variation of both groups are assumed to be 1.5.

This report shows the power for the indicated scenarios.

Plots Section



Non-Inferiority Tests for the Ratio of Two Means



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

Non-Inferiority Tests for the Ratio of Two Means

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. We will use the same settings as those given in Example 1. 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 and S1.

$$\begin{aligned}
 S1 &= \sqrt{\ln(COV^2 + 1)} \\
 &= \sqrt{\ln(1.5^2 + 1)} \\
 &= 1.085659 \\
 NIM' &= \ln(1 - NIM) \\
 &= \ln(0.8) \\
 &= 0.223144 \\
 D &= \ln(R1) \\
 &= \ln(0.95) \\
 &= -0.051293
 \end{aligned}$$

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for the Difference Between Two Means** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Difference Between Two Means**. You may then make the appropriate entries as listed below, or open **Example 2b** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Higher Means Are.....	Better
Nonparametric Adjustment.....	Ignore
Alpha	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group.....	100 to 1000 by 100
NIM (Non-Inferiority Margin)	0.223144
D (True Difference).....	-0.051293 0.0
S1 (Standard Deviation Group 1)	1.085659
S2 (Standard Deviation Group 2)	S1

Non-Inferiority Tests for the Ratio of Two Means

Output

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

Numeric Results**Numeric Results for Non-Inferiority Test (H0: Diff \leq -NIM; H1: Diff $>$ -NIM)**

Higher Means are Better

Test Statistic: T-Test

Power	N1	N2	N	-NIM	D	S1	S2	Alpha
0.19875	100	100	200	-0.223	-0.0513	1.0857	1.0857	0.025
0.35165	200	200	400	-0.223	-0.0513	1.0857	1.0857	0.025
0.49026	300	300	600	-0.223	-0.0513	1.0857	1.0857	0.025
0.60984	400	400	800	-0.223	-0.0513	1.0857	1.0857	0.025
0.70637	500	500	1000	-0.223	-0.0513	1.0857	1.0857	0.025
0.78275	600	600	1200	-0.223	-0.0513	1.0857	1.0857	0.025
0.84160	700	700	1400	-0.223	-0.0513	1.0857	1.0857	0.025
0.88599	800	800	1600	-0.223	-0.0513	1.0857	1.0857	0.025
0.91886	900	900	1800	-0.223	-0.0513	1.0857	1.0857	0.025
0.94284	1000	1000	2000	-0.223	-0.0513	1.0857	1.0857	0.025
0.30375	100	100	200	-0.223	0.0000	1.0857	1.0857	0.025
0.53604	200	200	400	-0.223	0.0000	1.0857	1.0857	0.025
0.70997	300	300	600	-0.223	0.0000	1.0857	1.0857	0.025
0.82799	400	400	800	-0.223	0.0000	1.0857	1.0857	0.025
0.90132	500	500	1000	-0.223	0.0000	1.0857	1.0857	0.025
0.94511	600	600	1200	-0.223	0.0000	1.0857	1.0857	0.025
0.97024	700	700	1400	-0.223	0.0000	1.0857	1.0857	0.025
0.98421	800	800	1600	-0.223	0.0000	1.0857	1.0857	0.025
0.99178	900	900	1800	-0.223	0.0000	1.0857	1.0857	0.025
0.99579	1000	1000	2000	-0.223	0.0000	1.0857	1.0857	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.