

Chapter 522

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Introduction

This procedure computes power and sample size for *non-inferiority* tests in two-sample designs in which the outcome is a continuous normal random variable and the variances of the two groups (populations) are assumed to be unequal. Measurements are made on individuals that have been randomly assigned to one of two groups. This is sometimes referred to as a *parallel-groups* design. This design is used in situations such as the comparison of the income level of two regions, the nitrogen content of two lakes, or the effectiveness of two drugs.

The details of sample size calculation for the two-sample design are presented in the Two-Sample T-Tests Allowing Unequal Variance chapter and they will not be duplicated here. This chapter only discusses those changes necessary for non-inferiority tests. Sample size formulas for non-inferiority tests of two means are presented in Chow et al. (2018) pages 50-51.

The Statistical Hypotheses

Remember that in the usual *t*-test setting, the null (H_0) and alternative (H_1) hypotheses for one-sided tests are defined as

$$H_0: \mu_1 - \mu_2 \leq \delta_0 \quad \text{versus} \quad H_1: \mu_1 - \mu_2 > \delta_0$$

or equivalently

$$H_0: \delta \leq \delta_0 \quad \text{versus} \quad H_1: \delta > \delta_0.$$

Rejecting this test implies that the mean difference is larger than the value δ_0 . This test is called an *upper-tailed test* because it is rejected in samples in which the difference between the sample means is larger than δ_0 .

Following is an example of a *lower-tailed test*.

$$H_0: \mu_1 - \mu_2 \geq \delta_0 \quad \text{versus} \quad H_1: \mu_1 - \mu_2 < \delta_0$$

or equivalently

$$H_0: \delta \geq \delta_0 \quad \text{versus} \quad H_1: \delta < \delta_0.$$

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Non-inferiority tests are special cases of the above directional tests. 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_1	Not used	<i>Mean</i> of population 1. Population 1 is assumed to consist of those who have received the new treatment.
μ_2	Not used	<i>Mean</i> of population 2. Population 2 is assumed to consist of those who have received the reference treatment.
M_{NI}	NIM	<i>Margin of non-inferiority</i> . This is a tolerance value that defines the magnitude of the amount that is not of practical importance. This may be thought of as the largest change from the baseline that is considered to be trivial. The sign of the value will be determined by the specific design that is being used.
δ	δ	<i>Actual difference</i> . This is the value of $\mu_1 - \mu_2$, the difference between the means. This is the value at which the power is calculated.

Note that the actual values of μ_1 and μ_2 are not needed. Only their difference is needed for power and sample size calculations.

Non-Inferiority Tests

A *non-inferiority test* tests that the treatment mean is not worse than the reference mean by more than the non-inferiority margin. The actual direction of the hypothesis depends on the response variable being studied.

Case 1: High Values Good

In this case, higher values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment mean is no less than a small amount below the reference mean. The value of δ at which power is calculated is often set to zero. The null and alternative hypotheses with $\delta_0 = -|M_{NI}|$ are

$$\begin{aligned} H_0: \mu_1 &\leq \mu_2 - |M_{NI}| && \text{versus} && H_1: \mu_1 > \mu_2 - |M_{NI}| \\ H_0: \mu_1 - \mu_2 &\leq -|M_{NI}| && \text{versus} && H_1: \mu_1 - \mu_2 > -|M_{NI}| \\ H_0: \delta &\leq -|M_{NI}| && \text{versus} && H_1: \delta > -|M_{NI}| \end{aligned}$$

Case 2: High Values Bad

In this case, lower values are better. The hypotheses are arranged so that rejecting the null hypothesis implies that the treatment mean is no more than a small amount above the reference mean. The value of δ at which power is calculated is often set to zero. The null and alternative hypotheses with $\delta_0 = |M_{NI}|$ are

$$\begin{aligned} H_0: \mu_1 &\geq \mu_2 + |M_{NI}| && \text{versus} && H_1: \mu_1 < \mu_2 + |M_{NI}| \\ H_0: \mu_1 - \mu_2 &\geq |M_{NI}| && \text{versus} && H_1: \mu_1 - \mu_2 < |M_{NI}| \\ H_0: \delta &\geq |M_{NI}| && \text{versus} && H_1: \delta < |M_{NI}| \end{aligned}$$

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Example

A non-inferiority test example will set the stage for the discussion of the terminology that follows. Suppose that a test is to be conducted to determine if a new cancer treatment adversely affects mean bone density. The adjusted mean bone density (AMBD) in the population of interest is 0.002300 gm/cm with a standard deviation of 0.000300 gm/cm. Clinicians decide that if the treatment reduces AMBD by more than 5% (0.000115 gm/cm), it poses a significant health threat. The treatment group standard deviation is 0.000350 gm/cm.

The hypothesis of interest is whether the mean AMBD in the treated group is more than 0.000115 below that of the reference group. The statistical test will be set up so that if the null hypothesis is rejected, the conclusion will be that the new treatment is non-inferior. The value 0.000115 gm/cm is called the *margin of non-inferiority*.

Two-Sample Unequal-Variance T-Test (Welch's T-Test) Statistic

Welch (1938) proposed the following test when the two variances are not assumed to be equal.

$$t_{df}^* = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_0}{s_{\bar{X}_1 - \bar{X}_2}^*}$$

where

$$\bar{X}_k = \frac{\sum_{i=1}^{n_k} X_{ki}}{n_k},$$

$$s_k = \sqrt{\left(\frac{\sum_{i=1}^{n_k} (X_{ki} - \bar{X}_k)^2}{(n_k - 1)} \right)},$$

$$s_{\bar{X}_1 - \bar{X}_2}^* = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

$$df = \frac{s_{\bar{X}_1 - \bar{X}_2}^{*4}}{\frac{s_1^4}{n_1^2(n_1 - 1)} + \frac{s_2^4}{n_2^2(n_2 - 1)}},$$

and δ_0 is the value of the difference hypothesized by the null hypothesis which depends on the magnitude and sign of M_{NI} .

The null hypothesis is rejected if the computed p-value is less than a specified level (usually 0.05). Otherwise, no conclusion can be reached.

Computing the Power

When $\sigma_1 \neq \sigma_2$, the power for Welch's unequal-variance t -test is calculated as follows.

1. Calculate: $\sigma_{\bar{x}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$.

2. Calculate: $df = \frac{\sigma_{\bar{x}}^4}{\frac{\sigma_1^4}{n_1^2(n_1-1)} + \frac{\sigma_2^4}{n_2^2(n_2-1)}}$

which is the adjusted degrees of freedom.

3. Find t_α such that $1 - T_{df}(t_\alpha) = \alpha$, where $T_{df}(x)$ is the area to the left of x under a central- t distribution with degrees of freedom, df .

4. Calculate the noncentrality parameter: $\lambda = \frac{\delta - \delta_0}{\sigma_{\bar{x}}}$.

5. Calculate: $Power = 1 - T'_{df,\lambda}(t_\alpha)$, where $T'_{df,\lambda}(x)$ is the area to the left of x under a noncentral- t distribution with degrees of freedom, df , and noncentrality parameter, λ .

When solving for something other than power, PASS uses this same power calculation formulation, but performs a search to determine that parameter.

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 most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be calculated from the values of the other parameters.

Select *Sample Size* when you want to determine the sample size needed to achieve a given power and alpha.

Select *Power* when you want to calculate the power of an experiment that has already been run.

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.

- **Better (H1: $\delta > -NIM$)**

If higher means are Better, the null hypothesis is $H_0: \delta \leq -NIM$, and the alternative hypothesis is $H_1: \delta > -NIM$.

- **Worse (H1: $\delta < NIM$)**

If higher means are Worse, the null hypothesis is $H_0: \delta \geq NIM$, and the alternative hypothesis is $H_1: \delta < NIM$.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

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 the null hypothesis should be rejected.

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 you reject the null hypothesis of inferiority when in fact the mean is not non-inferior.

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.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

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 $N - N1$.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

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.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Effect Size – Mean Difference

NIM (Non-Inferiority Margin)

This is the magnitude of the margin of non-inferiority. It must be entered as a positive number. If a negative value is entered, the absolute value is used.

When higher means are better, this value is the distance below the reference mean that is still considered non-inferior. When higher means are worse, this value is the distance above the reference mean that is still considered non-inferior.

δ (Actual Difference to Detect)

This is the actual difference between the treatment mean and the reference mean at which the power is calculated. For non-inferiority tests, this value is often set to zero. When higher means are better, $\delta > -\text{NIM}$. When higher means are worse, $\delta < \text{NIM}$.

Effect Size – Standard Deviations

σ_1 (Standard Deviation of Group 1)

The standard deviation entered here is the assumed standard deviation for the Group 1 population. σ_1 must be a positive number.

When σ_1 is not known, you must supply an estimate. Press the small ' σ ' button to the right to obtain calculation options for estimating the standard deviation.

σ_2 (Standard Deviation of Group 2)

The standard deviation entered here is the assumed standard deviation for the Group 2 population. σ_2 must be a positive number.

When σ_2 is not known, you must supply an estimate. Press the small ' σ ' button to the right to obtain calculation options for estimating the standard deviation.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Example 1 – Power Analysis

Suppose that a test is to be conducted to determine if a new cancer treatment adversely affects bone density. The adjusted mean bone density (AMBD) in the population of interest is 0.002300 gm/cm with a standard deviation of 0.000300 gm/cm. Clinicians decide that if the treatment reduces AMBD by more than 5% (0.000115 gm/cm), it poses a significant health threat. They also want to consider what would happen if the margin of equivalence is set to 2.5% (0.0000575 gm/cm). The treatment group standard deviation is 0.000350 gm/cm.

Following accepted procedure, the analysis will be a non-inferiority test using the t -test at the 0.025 significance level. Power is to be calculated assuming that the new treatment has no effect on AMBD. Several sample sizes between 10 and 800 will be analyzed. The researchers want to achieve a power of at least 90%. All numbers have been multiplied by 10000 to make the reports and plots easier to read.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Inferiority**, and then clicking on **Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance**. 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 (H1: $\delta > -NIM$)
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	10 50 100 200 300 500 600 800
NIM (Non-Inferiority Margin)	0.575 1.15
δ (Actual Difference to Detect)	0
σ_1 (Standard Deviation of Group 1)	3
σ_2 (Standard Deviation of Group 2)	3.5

Annotated Output

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

Numeric Results and Plots

Numeric Results for an Unequal-Variance T-Test

$$\delta = \mu_1 - \mu_2 = \mu_T - \mu_R$$

Higher Means are Better

Hypotheses: $H_0: \delta \leq -NIM$ vs. $H_1: \delta > -NIM$

Power	N1	N2	N	-NIM	δ	σ_1	σ_2	Alpha
0.05631	10	10	20	-0.575	0.0	3.0	3.5	0.025
0.13857	50	50	100	-0.575	0.0	3.0	3.5	0.025
0.23613	100	100	200	-0.575	0.0	3.0	3.5	0.025
0.42062	200	200	400	-0.575	0.0	3.0	3.5	0.025
0.57807	300	300	600	-0.575	0.0	3.0	3.5	0.025
0.79641	500	500	1000	-0.575	0.0	3.0	3.5	0.025
0.86323	600	600	1200	-0.575	0.0	3.0	3.5	0.025

(report continues)

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

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.
 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.

N1 and N2 are the number of items sampled from each population.

$N = N1 + N2$ is the total sample size.

-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 the reference mean that is still considered non-inferior.

$\delta = \mu1 - \mu2 = \muT - \muR$ is the difference between the treatment and reference means at which power and sample size calculations are made.

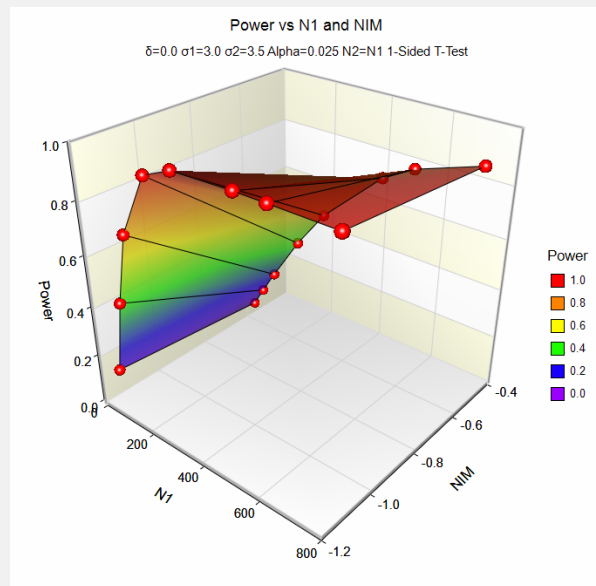
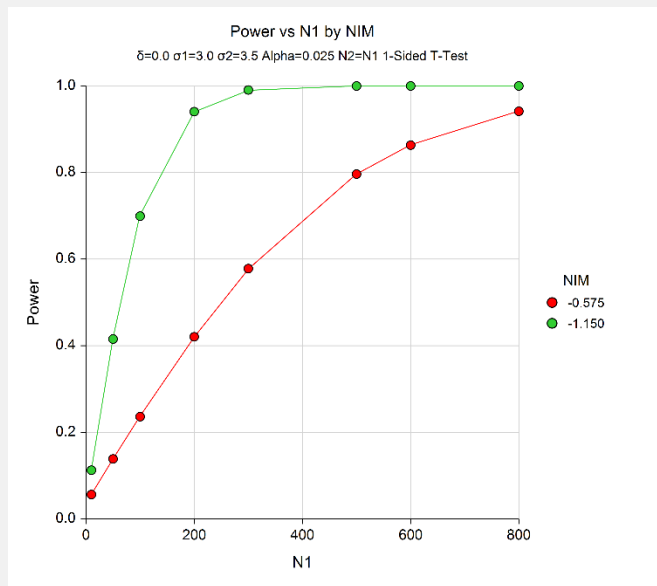
$\sigma1$ and $\sigma2$ are the assumed population standard deviations for groups 1 and 2, respectively.

Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

Group sample sizes of 10 and 10 achieve 6% power to detect non-inferiority using a one-sided, two-sample unequal-variance t-test. The margin of non-inferiority is -0.575. The actual difference between the means is assumed to be 0.0. The significance level (alpha) of the test is 0.025. The data are drawn from populations with standard deviations of 3.0 for group 1 and 3.5 for group 2.

Chart Section



The above report shows that for NIM = 1.15, the sample size necessary to obtain 90% power is about 170 per group. However, if NIM = 0.575, the required sample size is about 675 per group.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Example 2 – Finding the Sample Size

Continuing with Example 1, the researchers want to know the exact sample size for each value of NIM to achieve 90% power.

Setup

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

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Higher Means Are	Better (H1: $\delta > -NIM$)
Power	0.90
Alpha	0.025
Group Allocation	Equal (N1 = N2)
NIM (Non-Inferiority Margin)	0.575 1.15
δ (Actual Difference to Detect)	0
σ_1 (Standard Deviation of Group 1)	3
σ_2 (Standard Deviation of Group 2)	3.5

Output

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

Numeric Results

Numeric Results for an Unequal-Variance T-Test

$$\delta = \mu_1 - \mu_2 = \mu_T - \mu_R$$

Higher Means are Better

Hypotheses: $H_0: \delta \leq -NIM$ vs. $H_1: \delta > -NIM$

Target Power	Actual Power	N1	N2	N	-NIM	δ	σ_1	σ_2	Alpha
0.90	0.90018	676	676	1352	-0.575	0.0	3.0	3.5	0.025
0.90	0.90030	170	170	340	-1.150	0.0	3.0	3.5	0.025

This report shows the exact sample size requirement for each value of NIM.

Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance

Example 3 – Validation of Sample Size using Simulation

Example 5 of the Tests for Two Means (Simulation) procedure documentation presents a sample size calculation for a non-inferiority test of

$$H_0: \mu_1 - \mu_2 \geq 4 \quad \text{versus} \quad H_1: \mu_1 - \mu_2 < 4$$

or equivalently

$$H_0: \delta \geq 4 \quad \text{versus} \quad H_1: \delta < 4.$$

This example demonstrates an equal-variance t -test with standard deviation in both groups equal to 6, a NIM of 4, and an actual difference of 0. If we modify this example slightly such that the standard deviation in group 1 is 6 and the standard deviation in group 2 is 9, and we use the Welch's unequal-variance t -test (see Two Means (Simulation) – Example 8), then the sample size result is 78 per group with 10,000 simulations.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance** procedure window by expanding **Means**, then **Two Independent Means**, then clicking on **Non-Inferiority**, and then clicking on **Two-Sample T-Tests for Non-Inferiority Allowing Unequal Variance**. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Higher Means Are	Worse (H1: $\delta < \text{NIM}$)
Power	0.90
Alpha	0.025
Group Allocation	Equal (N1 = N2)
NIM (Non-Inferiority Margin)	4
δ (Actual Difference to Detect)	0
σ_1 (Standard Deviation of Group 1)	6
σ_2 (Standard Deviation of Group 2)	9

Output

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

Numeric Results

Numeric Results for an Unequal-Variance T-Test									
$\delta = \mu_1 - \mu_2 = \mu_T - \mu_R$									
Higher Means are Worse									
Hypotheses: $H_0: \delta \geq \text{NIM}$ vs. $H_1: \delta < \text{NIM}$									
Target Power	Actual Power	N1	N2	N	NIM	δ	σ_1	σ_2	Alpha
0.90	0.90018	78	78	156	4.0	0.0	6.0	9.0	0.025

The sample size of 78 per group matches the simulation result.