Chapter 207

Two-Sample T-Test from Means and SD's

Introduction

This procedure computes the two-sample t-test and several other two-sample tests directly from the mean, standard deviation, and sample size. Confidence intervals for the means, mean difference, and standard deviations can also be computed. Hypothesis tests included in this procedure can be produced for both one- and two-sided tests as well as equivalence tests.

Technical Details

The technical details and formulas for the methods of this procedure are presented in line with the Example 1 output. The output and technical details are presented in the following order:

- Confidence intervals of each group
- Confidence interval of $\mu_1 - \mu_2$
- Confidence intervals for individual group standard deviations
- Confidence interval of the standard deviation ratio
- Equal-Variance T-Test and associated power report
- Aspin-Welch Unequal-Variance T-Test and associated power report
- Z-Test
- Equivalence Test
- Variance Ratio Test

Null and Alternative Hypotheses

For comparing two means, the basic null hypothesis is that the means are equal,

$$H_0: \mu_1 = \mu_2$$

with three common alternative hypotheses,

$$H_a: \mu_1 \neq \mu_2 \text{ ,}$$
$$H_a: \mu_1 < \mu_2 \text{ , or}$$
$$H_a: \mu_1 > \mu_2 \text{ ,}$$

one of which is chosen according to the nature of the experiment or study.

A slightly different set of null and alternative hypotheses are used if the goal of the test is to determine whether μ_1 or μ_2 is greater than or less than the other by a given amount.

The null hypothesis then takes on the form

$$H_0: \mu_1 - \mu_2 = \textit{Hypothesized Difference}$$

and the alternative hypotheses,

$$H_a: \mu_1 - \mu_2 \neq \textit{Hypothesized Difference}$$
$$H_a: \mu_1 - \mu_2 < \textit{Hypothesized Difference}$$
$$H_a: \mu_1 - \mu_2 > \textit{Hypothesized Difference}$$

These hypotheses are equivalent to the previous set if the *Hypothesized Difference* is zero.

Data Structure

This procedure does not use data from the spreadsheet. Instead, you enter the values directly into the panel of the Data tab.

Example 1 – Analyzing Summarized Data

This section presents an example of using this panel to analyze a set of previously summarized data. A published report showed the following results: N1 = 15, Mean1 = 3.7122, SD1 = 1.9243, N2 = 13, Mean2 = 1.8934, and SD2 = 2.4531. Along with the other results, suppose you want to see an equivalence test in which the margin of equivalence is 0.3.

Setup

To run this example, complete the following steps:

1   Specify the Two-Sample T-Test from Means and SD's procedure options

- Find and open the **Two-Sample T-Test from Means and SD's** procedure using the menus or the Procedure Navigator.
- The settings 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.

Data Tab

Group 1 Sample Size..**15**
Group 1 Mean..**3.7122**
Group 1 Standard Deviation **1.9243**
Group 2 Sample Size..**13**
Group 2 Mean..**1.8934**
Group 2 Standard Deviation **2.4531**

Reports Tab

Confidence Level...**95**
Confidence Intervals of Each Group Mean ......**Checked**
Confidence Interval of μ1 - μ2.........................**Checked**
  Limits ....**Two-Sided**
Confidence Intervals of Each Group SD ....**Checked**
Confidence Interval of σ1/σ2 ....**Checked**
Alpha **0.05**
H0 μ1 - μ2 =...**0.0**
Ha ....**Two-Sided and One-Sided** (Usually, a single alternative hypothesis is chosen, but all three alternatives are shown in this example to exhibit all the reporting options.)
Equal-Variance T-Test....................................**Checked**
Unequal-Variance T-Test................................**Checked**
Z-Test ....**Checked**
Equivalence Test ....**Checked**
  Upper Equivalence Bound **0.3**
  Lower Equivalence Bound ....**-0.3**
Power Report for Equal-Variance T-Test.........**Checked**
Power Report for Unequal-Variance T-Test.....**Checked**
Variance-Ratio Test ....**Checked**

2   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Confidence Intervals for the Group Means (µ's)

Confidence Intervals for the Group Means (µ's)

Group	N	Mean	Standard Deviation of the Data	Standard Error of the Mean	T*	95% Confidence Interval Limits for the Mean (µ)	
						Lower	Upper
1	15	3.7122	1.9243	0.4968521	2.1448	2.646558	4.777842
2	13	1.8934	2.4531	0.6803675	2.1788	0.4110065	3.375793

This report documents the values that were input alone with the associated confidence intervals.

N

This is the number of individuals in the group, or the group sample size.

Mean

This is the average for each group.

Standard Deviation of the Data

The sample standard deviation is the square root of the sample variance. It is a measure of spread.

Standard Error of the Mean

This is the estimated standard deviation for the distribution of sample means for an infinite population. It is the sample standard deviation divided by the square root of sample size, n.

T*

This is the t-value used to construct the confidence interval. If you were constructing the interval manually, you would obtain this value from a table of the Student's t distribution with n - 1 degrees of freedom.

95% Confidence Interval Limits for the Mean (µ) (Lower and Upper)

These are the lower and upper limits of an interval estimate of the mean based on a Student's t distribution with n - 1 degrees of freedom. This interval estimate assumes that the population standard deviation is not known and that the data are normally distributed.

Confidence Interval for the Mean Difference (μ1 - μ2)

Two-Sided Confidence Interval for the Mean Difference (μ1 - μ2)

Variance Assumption	DF	Mean Difference	Standard Error	T*	95% Confidence Interval Limits for the Mean Difference (μ1 - μ2)	
					Lower	Upper
Equal	26	1.8188	0.8277123	2.0555	0.117413	3.520187
Unequal	22.68	1.8188	0.8424737	2.0703	0.07465944	3.562941

This report provides confidence intervals for the difference between the means. The first row gives the equal-variance interval. The second row gives the interval based on the unequal-variance assumption formula.

The interpretation of these confidence intervals is that when populations are repeatedly sampled and confidence intervals are calculated, 95% of those confidence intervals will include (cover) the true value of the difference.

DF

The degrees of freedom are used to determine the T distribution from which T-alpha is generated.

For the equal variance case:

$$df = n_1 + n_2 - 2$$

For the unequal variance case:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Mean Difference

This is the difference between the sample means, $\overline{X}_1 - \overline{X}_2$.

Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

For the equal variance case:

$$SE_{\overline{X}_1-\overline{X}_2} = \sqrt{\left(\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

For the unequal variance case:

$$SE_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

T*

This is the t-value used to construct the confidence limits. It is based on the degrees of freedom and the confidence level.

95% Confidence Interval Limits for the Mean Difference (Lower and Upper)

These are the confidence limits of the confidence interval for $\mu_1 - \mu_2$. The confidence interval formula is

$$\overline{X}_1 - \overline{X}_2 \pm T_{df}^* SE_{\overline{X}_1-\overline{X}_2}$$

The equal-variance and unequal-variance assumption formulas differ by the values of T* and the standard error.

Confidence Intervals for the Group Standard Deviations (σ's)

Confidence Intervals for the Group Standard Deviations (σ's)

Group	N	Standard Deviation of the Data	95% Confidence Interval Limits for the Standard Deviation (σ)	
			Lower	Upper
1	15	1.9243	1.408831	3.034812
2	13	2.4531	1.759084	4.049418

This report gives a confidence interval for the standard deviation in each group. Note that the usefulness of these intervals is very dependent on the assumption that the data are sampled from a normal distribution.

Using the common notation for sample statistics (see, for example, ZAR (1984) page 115), a $100(1-\alpha)\%$ confidence interval for the standard deviation is given by

$$\sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha/2,n-1}^2}} \leq \sigma \leq \sqrt{\frac{(n-1)s^2}{\chi_{\alpha/2,n-1}^2}}$$

Confidence Interval for the Standard Deviation Ratio (σ1 / σ2)

Confidence Interval for the Standard Deviation Ratio (σ1 / σ2)

		Standard Deviation			95% Confidence Interval Limits for the Standard Deviation Ratio (σ1 / σ2)	
N1	N2	SD1	SD2	Ratio SD1 / SD2	Lower	Upper
15	13	1.9243	2.4531	0.784436	0.4380881	1.369993

This report gives a confidence interval for the ratio of the two standard deviations. Note that the usefulness of these intervals is very dependent on the assumption that the data are sampled from a normal distribution.

Using the common notation for sample statistics (see, for example, ZAR (1984) page 125), a $100(1 - \alpha)\%$ confidence interval for the ratio of two standard deviations is given by

$$\frac{s_1}{s_2\sqrt{F_{1-\alpha/2,n_1-1,n_2-1}}} \le \frac{\sigma_1}{\sigma_2} \le \frac{s_1\sqrt{F_{1-\alpha/2,n_2-1,n_1-1}}}{s_2}$$

Equal-Variance T-Test

This section presents the results of the traditional equal-variance T-test. Here, reports for all three alternative hypotheses are shown, but a researcher would typically choose one of the three before generating the output. All three tests are shown here for the purpose of exhibiting all the output options available.

Equal-Variance T-Test

Alternative Hypothesis	Mean Difference	Standard Error	T-Statistic	DF	P-Value	Reject H0 at α = 0.05?
μ1 - μ2 ≠ 0	1.8188	0.8277123	2.1974	26	0.03710	Yes
μ1 - μ2 < 0	1.8188	0.8277123	2.1974	26	0.98145	No
μ1 - μ2 > 0	1.8188	0.8277123	2.1974	26	0.01855	Yes

Alternative Hypothesis

The (unreported) null hypothesis is

$$H_0 : \mu_1 - \mu_2 = \textit{Hypothesized Difference} = 0$$

and the alternative hypotheses,

$$H_a : \mu_1 - \mu_2 \ne \textit{Hypothesized Difference} = 0$$

$$H_a : \mu_1 - \mu_2 < \textit{Hypothesized Difference} = 0$$

$$H_a : \mu_1 - \mu_2 > \textit{Hypothesized Difference} = 0$$

Since the *Hypothesized Difference* is zero in this example, the null and alternative hypotheses can be simplified to

Null hypothesis:

$$H_0 : \mu_1 = \mu_2$$

Alternative hypotheses:

$$H_a : \mu_1 \neq \mu_2 \ ,$$

$$H_a : \mu_1 < \mu_2 \ , \text{ or}$$

$$H_a : \mu_1 > \mu_2 \ .$$

In practice, the alternative hypothesis should be chosen in advance.

Mean Difference

This is the difference between the sample means, $\overline{X}_1 - \overline{X}_2$.

Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

The formula for the standard error of the difference in the equal-variance T-test is:

$$SE_{\overline{X}_1 - \overline{X}_2} = \sqrt{ \left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) }$$

T-Statistic

The T-Statistic is the value used to produce the *p*-value (Prob Level) based on the *T* distribution. The formula for the T-Statistic is:

$$T - Statistic = \frac{\overline{X}_1 - \overline{X}_2 - Hypothesized\ Difference}{SE_{\overline{X}_1 - \overline{X}_2}}$$

In this instance, the hypothesized difference is zero, so the T-Statistic formula reduces to

$$T - Statistic = \frac{\overline{X}_1 - \overline{X}_2}{SE_{\overline{X}_1 - \overline{X}_2}}$$

DF

The degrees of freedom define the T *distribution* upon which the probability values are based. The formula for the degrees of freedom in the equal-variance T-test is:

$$df = n_1 + n_2 - 2$$

P-Value

The *p*-value, also known as the probability level or significance level, is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the *p*-value is less than the prescribed $α$, in this case 0.05, the null hypothesis is rejected in favor of the alternative hypothesis. Otherwise, there is not sufficient evidence to reject the null hypothesis.

Reject H0 at α = 0.05?

This column indicates whether or not the null hypothesis is rejected, in favor of the alternative hypothesis, based on the *p*-value and chosen $α$. A test in which the null hypothesis is rejected is sometimes called *significant*.

Power for the Equal-Variance T-Test

The power report gives the power of a test where it is assumed that the population means and standard deviations are equal to the sample means and standard deviations. Powers are given for alpha values of 0.05 and 0.01. For a much more comprehensive and flexible investigation of power or sample size, we recommend you use the **PASS** software program.

Power for the Equal-Variance T-Test

Alternative Hypothesis	Sample Size		Mean		Standard Deviation		Power	
	N1	N2	μ1	μ2	σ1	σ2	α = 0.05	α = 0.01
μ1 - μ2 ≠ 0	15	13	3.7122	1.8934	1.9243	2.4531	0.56195	0.30252
μ1 - μ2 < 0	15	13	3.7122	1.8934	1.9243	2.4531	0.00008	0.00001
μ1 - μ2 > 0	15	13	3.7122	1.8934	1.9243	2.4531	0.68962	0.40403

This report assumes that the population means and standard deviations are equal to the sample values.

Alternative Hypothesis

This value identifies the test direction of the test reported in this row. In practice, you would select the alternative hypothesis prior to your analysis and have only one row showing here.

N1 and N2

N1 and N2 are the assumed sample sizes for groups 1 and 2.

μ1 and μ2

These are the assumed population means on which the power calculations are based.

σ1 and σ2

These are the assumed population standard deviations on which the power calculations are based.

Power (α = 0.05) and Power (α = 0.01)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of a type II error (β). The power of the test depends on the sample size, the magnitudes of the standard deviations, the alpha level, and the true difference between the two population means.

The power value calculated here assumes that the population standard deviations are equal to the sample standard deviations and that the difference between the population means is exactly equal to the difference between the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false.

Some ways to increase the power of a test include the following:

1. Increase the alpha level. Perhaps you could test at alpha = 0.05 instead of alpha = 0.01.

2. Increase the sample size.

3. Decrease the magnitude of the standard deviations. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

Aspin-Welch Unequal-Variance T-Test

This section presents the results of the traditional equal-variance T-test. Here, reports for all three alternative hypotheses are shown, but a researcher would typically choose one of the three before generating the output. All three tests are shown here for the purpose of exhibiting all the output options available.

Aspin-Welch Unequal-Variance T-Test

Alternative Hypothesis	Mean Difference	Standard Error	T-Statistic	DF	P-Value	Reject H0 at α = 0.05?
μ1 - μ2 ≠ 0	1.8188	0.8424737	2.1589	22.68	0.04169	Yes
μ1 - μ2 < 0	1.8188	0.8424737	2.1589	22.68	0.97916	No
μ1 - μ2 > 0	1.8188	0.8424737	2.1589	22.68	0.02084	Yes

Alternative Hypothesis

The (unreported) null hypothesis is

$$H_0: \mu_1 - \mu_2 = \text{Hypothesized Difference} = 0$$

and the alternative hypotheses,

$$H_a: \mu_1 - \mu_2 \neq \text{Hypothesized Difference} = 0$$

$$H_a: \mu_1 - \mu_2 < \text{Hypothesized Difference} = 0$$

$$H_a: \mu_1 - \mu_2 > \text{Hypothesized Difference} = 0$$

Since the *Hypothesized Difference* is zero in this example, the null and alternative hypotheses can be simplified to

Null hypothesis:

$$H_0: \mu_1 = \mu_2$$

Alternative hypotheses:

$$H_a: \mu_1 \neq \mu_2 \, ,$$

$$H_a: \mu_1 < \mu_2 \, , \text{ or}$$

$$H_a: \mu_1 > \mu_2 \, .$$

In practice, the alternative hypothesis should be chosen in advance.

Mean Difference

This is the difference between the sample means, $\overline{X}_1 - \overline{X}_2$.

Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

The formula for the standard error of the difference in the Aspin-Welch unequal-variance T-test is:

$$SE_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

T-Statistic

The T-Statistic is the value used to produce the *p*-value (Prob Level) based on the *T* distribution. The formula for the T-Statistic is:

$$T - Statistic = \frac{\overline{X}_1 - \overline{X}_2 - Hypothesized\ Difference}{SE_{\overline{X}_1 - \overline{X}_2}}$$

In this instance, the hypothesized difference is zero, so the T-Statistic formula reduces to

$$T - Statistic = \frac{\overline{X}_1 - \overline{X}_2}{SE_{\overline{X}_1 - \overline{X}_2}}$$

DF

The degrees of freedom define the T *distribution* upon which the probability values are based. The formula for the degrees of freedom in the Aspin-Welch unequal-variance T-test is:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

P-Value

The *p*-value, also known as the probability level or significance level, is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the *p*-value is less than the prescribed *α*, in this case 0.05, the null hypothesis is rejected in favor of the alternative hypothesis. Otherwise, there is not sufficient evidence to reject the null hypothesis.

Reject H0 at α = 0.05?

This column indicates whether or not the null hypothesis is rejected, in favor of the alternative hypothesis, based on the p-value and chosen *α*. A test in which the null hypothesis is rejected is sometimes called *significant*.

Power for the Aspin-Welch Unequal-Variance T-Test

The power report gives the power of a test where it is assumed that the population means and standard deviations are equal to the sample means and standard deviations. Powers are given for alpha values of 0.05 and 0.01. For a much more comprehensive and flexible investigation of power or sample size, we recommend you use the **PASS** software program.

Power for the Aspin-Welch Unequal-Variance T-Test

Alternative Hypothesis	Sample Size		Mean		Standard Deviation		Power	
	N1	N2	μ1	μ2	σ1	σ2	α = 0.05	α = 0.01
μ1 - μ2 ≠ 0	15	13	3.7122	1.8934	1.9243	2.4531	0.54266	0.28341
μ1 - μ2 < 0	15	13	3.7122	1.8934	1.9243	2.4531	0.00009	0.00001
μ1 - μ2 > 0	15	13	3.7122	1.8934	1.9243	2.4531	0.67326	0.38377

This report assumes that the population means and standard deviations are equal to the sample values.

Alternative Hypothesis

This value identifies the test direction of the test reported in this row. In practice, you would select the alternative hypothesis prior to your analysis and have only one row showing here.

N1 and N2

N1 and N2 are the assumed sample sizes for groups 1 and 2.

μ1 and μ2

These are the assumed population means on which the power calculations are based.

σ1 and σ2

These are the assumed population standard deviations on which the power calculations are based.

Power (α = 0.05) and Power (α = 0.01)

Power is the probability of rejecting the hypothesis that the means are equal when they are in fact not equal. Power is one minus the probability of a type II error (*β*). The power of the test depends on the sample size, the magnitudes of the standard deviations, the alpha level, and the true difference between the two population means.

The power value calculated here assumes that the population standard deviations are equal to the sample standard deviations and that the difference between the population means is exactly equal to the difference between the sample means.

High power is desirable. High power means that there is a high probability of rejecting the null hypothesis when the null hypothesis is false.

Some ways to increase the power of a test include the following:

1. Increase the alpha level. Perhaps you could test at alpha = 0.05 instead of alpha = 0.01.

2. Increase the sample size.

3. Decrease the magnitude of the standard deviations. Perhaps you can redesign your study so that measurements are more precise and extraneous sources of variation are removed.

Z-Test

This section presents the results of the two-sample Z-test, which is used when the population standard deviations are known. Because the population standard deviations are rarely known, this test is not commonly used in practice. The Z-test is included in this procedure for the less-common case of known standard deviations, and for two-sample hypothesis test training. In this example, reports for all three alternative hypotheses are shown, but a researcher would typically choose one of the three before generating the output. All three tests are shown here for the purpose of exhibiting all the output options available.

Z-Test

Alternative Hypothesis	Mean Difference	σ1	σ2	Standard Error	Z-Statistic	P-Value	Reject H0 at α = 0.05?
μ1 - μ2 ≠ 0	1.8188	1.9243	2.4531	0.8424737	2.1589	0.03086	Yes
μ1 - μ2 < 0	1.8188	1.9243	2.4531	0.8424737	2.1589	0.98457	No
μ1 - μ2 > 0	1.8188	1.9243	2.4531	0.8424737	2.1589	0.01543	Yes

This report assumes that the population standard deviations are equal to the sample values.

Alternative Hypothesis

The (unreported) null hypothesis is

$$H_0: \mu_1 - \mu_2 = \text{Hypothesized Difference} = 0$$

and the alternative hypotheses,

$$H_a: \mu_1 - \mu_2 \neq \text{Hypothesized Difference} = 0$$
$$H_a: \mu_1 - \mu_2 < \text{Hypothesized Difference} = 0$$
$$H_a: \mu_1 - \mu_2 > \text{Hypothesized Difference} = 0$$

Since the *Hypothesized Difference* is zero in this example, the null and alternative hypotheses can be simplified to

Null hypothesis:

$$H_0: \mu_1 = \mu_2$$

Alternative hypotheses:

$$H_a: \mu_1 \neq \mu_2 \,,$$
$$H_a: \mu_1 < \mu_2 \,, \text{ or}$$
$$H_a: \mu_1 > \mu_2 \,.$$

In practice, the alternative hypothesis should be chosen in advance.

Mean Difference

This is the difference between the sample means, $\overline{X}_1 - \overline{X}_2$.

σ1 and σ2

These are the known Group 1 and Group 2 population standard deviations.

Standard Error

This is the estimated standard deviation of the distribution of differences between independent sample means.

The formula for the standard error of the difference in the Z-test is:

$$SE_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Z-Statistic

The Z-Statistic is the value used to produce the *p*-value (Prob Level) based on the *Z* distribution. The formula for the Z-Statistic is:

$$Z - Statistic = \frac{\overline{X}_1 - \overline{X}_2 - Hypothesized\ Difference}{SE_{\overline{X}_1 - \overline{X}_2}}$$

In this instance, the hypothesized difference is zero, so the Z-Statistic formula reduces to

$$Z - Statistic = \frac{\overline{X}_1 - \overline{X}_2}{SE_{\overline{X}_1 - \overline{X}_2}}$$

P-Value

The *p*-value, also known as the probability level or significance level, is the probability that the test statistic will take a value at least as extreme as the observed value, assuming that the null hypothesis is true. If the *p*-value is less than the prescribed *α*, in this case 0.05, the null hypothesis is rejected in favor of the alternative hypothesis. Otherwise, there is not sufficient evidence to reject the null hypothesis.

Reject H0 at α = 0.05?

This column indicates whether or not the null hypothesis is rejected, in favor of the alternative hypothesis, based on the *p*-value and chosen *α*. A test in which the null hypothesis is rejected is sometimes called *significant*.

Equivalence Test of the Means

Equivalence Tests of the Means

Hypotheses: H0: μ1 - μ2 ≤ -0.3 or μ1 - μ2 ≥ 0.3 versus Ha: -0.3 < μ1 - μ2 < 0.3
Mean Difference: 1.8188

Variance Assumption	One-Sided Test #1 (H01: μ1 - μ2 ≤ -0.3 vs. Ha1: μ1 - μ2 > -0.3)		One-Sided Test #2 (H02: μ1 - μ2 ≥ 0.3 vs. Ha2: μ1 - μ2 < 0.3)		Overall Equivalence Test P-Value	Reject H0 and Conclude Equivalence at α = 0.05?
	Test Statistic	P-Value	Test Statistic	P-Value		
Equal	2.5598	0.00832	1.8349	0.96101	0.96101	No
Unequal	2.5150	0.00973	1.8028	0.95764	0.95764	No

This report gives the results the equivalence test. The equivalence test is designed to establish that, for practical purposes, the two means are equal. To accomplish this, the roles of the null and alternative hypotheses are reversed. The hypotheses for testing equivalence of two means are (assuming that $\delta_L < 0$ and $\delta_U > 0$)

$$H_0: \mu_1 - \mu_2 \leq \delta_L \text{ or } \mu_1 - \mu_2 \geq \delta_U \text{ versus } H_a: \delta_L < \mu_1 - \mu_2 < \delta_U$$

The alternative hypothesis states that the true difference is in some small, clinically acceptable range. For example, we might be willing to conclude that the benefits of two drugs are equivalent if the difference in their mean response rates is between -0.1 and 0.1.

The conventional method of testing equivalence hypotheses is to perform two, one-sided tests (TOST) of hypotheses. The null hypothesis of non-equivalence is rejected in favor of the alternative hypothesis of equivalence if both one-sided tests are rejected. Unlike the common two-sided tests, however, the type I error rate for each test is set directly at the nominal level (usually 0.05)—it is not split in half. So, to perform the test, two one-sided tests are conducted at the α significance level. If both are rejected, the alternative hypothesis is concluded at the α significance level. The p-value of the overall equivalence test is the maximum of the p-values of the two individual tests.

The two, one-sided tests of hypotheses for the difference are

$$H_{01}: \mu_1 - \mu_2 \leq \delta_L \text{ versus } H_{a1}: \mu_1 - \mu_2 > \delta_L$$

$$H_{02}: \mu_1 - \mu_2 \geq \delta_U \text{ versus } H_{a2}: \mu_1 - \mu_2 < \delta_U$$

H01 and Ha1

The Test Statistic is the T-statistic for testing the first one-sided hypothesis. The Prob Level is the p-value for this test. The values are reported for both the equal-variance formula and the unequal-variance formula.

H02 and Ha2

The Test Statistic is the T-statistic for testing the second one-sided hypothesis. The Prob Level is the p-value for this test. The values are reported for both the equal-variance formula and the unequal-variance formula.

Overall Equivalence Test P-Value

The Overall p-value is given as the larger of the two one-sided test p-values. This p-value corresponds to the overall hypotheses of non-equivalence versus equivalence.

Reject H0 and Conclude Equivalence at α = 0.05?

By comparing the overall p-value to the α-level, a yes or no answer to this conclusion is given.

Variance Ratio Test

Variance-Ratio Test

Degrees of Freedom				Reject Hypothesis of Equal Variances at α = 0.05?
DF1	DF2	F-Value	P-Value	
14	12	1.6251	0.38335	No

This report provides a test of the null hypothesis of equal variances. It relies heavily on the assumption that the data of each group are sampled from a normal distribution. Failing to reject equality does not imply equality.

Degrees of Freedom (DF1 and DF2)

These are the degrees of freedom of the numerator and the denominator of the test. Each value is the group sample size minus one.

F-Value

The F-Value is the F-statistic for the test. The formula is the ratio of the larger variance to the smaller variance:

$$F = \frac{larger\ s^2}{smaller\ s^2}$$

P-Value

This is the associated *p*-value for the variance ratio test.

Reject Hypothesis of Equal Variances at α = 0.05?

By comparing the overall *p*-value to the *α*-level, a yes or no conclusion is given.