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

Data Structure

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

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,

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

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

$$H_0: \mu_1 - \mu_2 = Hypothesized\ Difference$$

and the alternative hypotheses,

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

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

Procedure Options

This section describes the options available in this procedure.

Data Tab

Enter the data values directly on this panel.

Group Summary Values

Sample Sizes

These boxes are used to specify the sample sizes (number of subjects) in each group.

Means

These boxes are used to specify the sample means for each group. For power calculations, these means are the assumed population means.

Standard Deviations

These boxes are used to specify the standard deviation for each group. For the Z-test and the power calculations, these standard deviations are assumed to be the population standard deviations. For all other tests and confidence intervals, these are the sample standard deviations.

Reports Tab

The options on this panel specify which reports will be included in the output.

Confidence Intervals

Confidence Level

This confidence level is used for all confidence interval reports below. The confidence level reflects the percent of confidence that the true mean difference (or mean, standard deviation, or ratio) is between the confidence interval limits. Typical confidence levels are 90%, 95%, and 99%, with 95% being the most common.

Confidence Intervals of Each Group Mean

This section reports the sample size, mean, standard deviation, standard error, and confidence interval of the mean, for each group.

Confidence Interval of $\mu_1 - \mu_2$

This section reports the confidence interval for the difference between the two means for both the equal variance and unequal variance assumption formulas.

Limits

Specify whether a two-sided or one-sided confidence interval of the difference ($\mu_1 - \mu_2$) is to be reported.

- **Two-Sided**

  For this selection, the lower and upper limits of $\mu_1 - \mu_2$ are reported, giving a confidence interval of the form (Lower Limit, Upper Limit).

- **One-Sided Upper**

  For this selection, only an upper limit of $\mu_1 - \mu_2$ is reported, giving a confidence interval of the form (-∞, Upper Limit).

- **One-Sided Lower**

  For this selection, only a lower limit of $\mu_1 - \mu_2$ is reported, giving a confidence interval of the form (Lower Limit, ∞).

Confidence Intervals of Each Group Standard Deviation

This section reports two-sided confidence intervals for the standard deviation of each group individually. This standard deviation confidence interval formula relies heavily on the normal distribution assumption.

Confidence Interval of σ1 / σ2

This section reports a two-sided confidence interval for the ratio of the standard deviations. This standard deviation ratio confidence interval formula relies heavily on the normal distribution assumption.

Tests

Alpha

Alpha is the significance level used in the hypothesis tests. A value of 0.05 is most commonly used, but 0.1, 0.025, 0.01, and other values are sometimes used. Typical values range from 0.001 to 0.20.

H0 Value

This is the hypothesized difference between the two population means. To test only whether there is a difference between the two groups, this value is set to zero. If, for example, a researcher wished to test whether the Group 1 population mean is greater than the Group 2 population mean by at least 5, this value would be set to 5.

Alternative Hypothesis Direction

Specify the direction of the alternative hypothesis. This direction applies to the three mean comparison tests. The selection 'Two-Sided and One-Sided' will produce all three tests for each mean comparison test selected.

Tests - Parametric

Equal Variance T-Test

This section reports the details of the T-Test for comparing two means when the variance of the two groups is assumed to be equal.

Power Report for Equal Variance T-Test

This report gives the power of the equal variance T-Test when it is assumed that the entered values are the population mean and population standard deviation.

Unequal Variance T-Test

This section reports the details of the Aspin-Welch T-Test method for comparing two means when the variance of the two groups is assumed to be unequal.

Power Report for Unequal Variance T-Test

This report gives the power of the Aspin-Welch unequal variance T-Test when it is assumed that the entered values are the population mean and population standard deviation.

Z-Test

This report gives the Z-Test for comparing two means and assumes that the standard deviations entered are the known population standard deviations. This test is rarely used in practice because the population standard deviations are rarely known.

Equivalence Test

The equivalence test report gives the results of the two one-sided tests for determining equivalence. The results are reported for both the equal-variance assumption formulas and the unequal-variance assumption formulas. The null hypothesis for this test is that the two means are different by some pre-specified amount. The alternative hypothesis is that the difference in means is inside the pre-specified bounds. The *p*-value for this test is the larger of the two one-sided test *p*-values.

Upper Equivalence Bound

Specify the upper bound for the equivalence test of the difference in means (Mean 1 – Mean 2). If the difference in the two population means is greater than this value, the two groups are not "equivalent." This value is sometimes called the "Margin of Equivalence.'

Lower Equivalence Bound

Specify the lower bound for the equivalence test of the difference in means (Mean 1 – Mean 2). If the difference in the two population means is less than this value, the two groups are not "equivalent." This value is sometimes called the "Margin of Equivalence." If this value is left blank, then the negative of the Upper Equivalence Bound is used.

Variance Ratio Test

This report gives the results of the variance ratio test. The variance ratio test is used to determine whether there is sufficient evidence to claim that the variances of the two groups differ. This variance ratio test depends heavily on the assumption that the two samples come from normal distributions.

Decimal Places

Decimal Places

This option allows the user to specify the number of decimal places directly or based on the significant digits. If one of the significant digits options is used, the ending zero digits are not shown. For example, if Auto (Up to 7)' is chosen,

  0.0500 is displayed as 0.05
  1.314583689 is displayed as 1.314584

The output formatting system is not designed to accommodate 'Auto (Up to 13)', and if chosen, this will likely lead to lines that run on to a second line. This option is included, however, for the rare case when a very large number of decimals is needed.

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.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Two-Sample T-Test from Means and SD's window.

1   Open the Two-Sample T-Test from Means and SD's window.
- On the menus, select **Analysis**, then **T-Tests**, then **Two-Sample T-Test from Means and SD's**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

2   Specify the data.
- Select the **Data tab**.
- In the **Sample Size Group 1** box, enter **15**.
- In the **Mean Group 1** box, enter **3.7122**.
- In the **Standard Deviation Group 1** box, enter **1.9243**.
- In the **Sample Size Group 2** box, enter **13**.
- In the **Mean Group 2** box, enter **1.8934**.
- In the **Standard Deviation Group 2** box, enter **2.4531**.

3   Specify the reports.
- Select the **Reports tab**.
- Leave the Confidence Level at 95%.
- Make sure the Confidence Intervals of Each Group Mean check box is checked.
- Make sure the Confidence Interval of $\mu1 – \mu2$ check box is checked.
- Leave the Confidence Interval **Limits** at **Two-Sided**.
- Make sure the Confidence Intervals of Each Group Standard Deviation check box is checked.
- Make sure the Confidence Interval of $\sigma1/\sigma2$ check box is checked.
- Leave **Alpha** at **0.05**.
- Leave the **H0 Value** at **0.0**.
- For **Ha,** select **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.
- Make sure the **Equal Variance T-Test** check box is checked.
- Check the Power Report for Equal Variance T-Test check box.

- Make sure the **Unequal Variance T-Test** check box is checked.
- Check the Power Report for Unequal Variance T-Test check box.
- Check the **Z-Test** check box. Enter **150** for σ1, and **110** for σ2.
- Check the **Equivalence Test** check box. Enter **0.3** for the Upper Equivalence Bound and **-0.3** for the Lower Equivalence Bound.
- Make sure the **Variance Ratio Test** check box is checked.

4   Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

The following reports will be displayed in the Output window.

Confidence Intervals of Means

Confidence Intervals of Means

					95.0% C. I. of μ	
Group	N	Mean	Standard Deviation	Standard Error	Lower Limit	Upper Limit
1	15	3.7122	1.9243	0.4968521	2.646558	4.777842
2	13	1.8934	2.4531	0.6803675	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

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

Standard Error

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*.

Confidence Interval Lower Limit

This is the lower limit 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 Upper Limit

This is the upper limit of the interval estimate for the mean based on a *t* distribution with *n* - 1 degrees of freedom.

Confidence Intervals of Difference

Two-Sided Confidence Interval for μ1 – μ2

						95.0% C. I. of μ1 - μ2	
Variance Assumption	DF	Mean Difference	Standard Deviation	Standard Error	T*	Lower Limit	Upper Limit
Equal	26	1.8188	2.184327	0.8277123	2.0555	0.117413	3.520187
Unequal	22.68	1.8188	3.117793	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 Deviation

In the equal variance case, this quantity is:

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

In the unequal variance case, this quantity is:

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{s_1^2 + s_2^2}$$

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.

Lower and Upper Confidence Limits

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

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

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

Confidence Intervals of Standard Deviations

Confidence Intervals of Standard Deviations

Sample	N	Mean	Standard Deviation	Standard Error	95.0% C. I. of σ Lower Limit	Upper Limit
1	15	3.7122	1.9243	0.4968521	1.408831	3.034812
2	13	1.8934	2.4531	0.6803675	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

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}} \le \sigma \le \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}}$$

Confidence Interval for Standard Deviation Ratio

Confidence Interval for Standard Deviation Ratio (σ1 / σ2)

N1	N2	SD1	SD2	SD1 / SD2	95.0% C. I. of σ1 / σ2 Lower Limit	Upper Limit
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 Section

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 of Difference	T-Statistic	d.f.	Prob Level	Reject H0 at α = 0.050
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

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

and the alternative hypotheses,

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

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

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

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 of Difference

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

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

d.f.

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:

$$df = n_1 + n_2 - 2$$

Prob Level

The probability level, also known as the *p*-value 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 = (0.050)

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

Alternative Hypothesis	N1	N2	$\mu 1$	$\mu 2$	$\sigma 1$	$\sigma 2$	Power (α = 0.05)	Power (α = 0.01)
$\mu 1 - \mu 2 \neq 0$	15	13	3.7122	1.8934	1.9243	2.4531	0.56195	0.30252
$\mu 1 - \mu 2 < 0$	15	13	3.7122	1.8934	1.9243	2.4531	0.00008	0.00001
$\mu 1 - \mu 2 > 0$	15	13	3.7122	1.8934	1.9243	2.4531	0.68962	0.40403

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.

$\mu 1$ and $\mu 2$

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

$\sigma 1$ and $\sigma 2$

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 =.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 Section

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 of Difference	T-Statistic	d.f.	Prob Level	Reject H0 at α = 0.050
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

$$H_0: \mu_1 - \mu_2 = Hypothesized\ Difference = 0$$

and the alternative hypotheses,

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

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

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

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:

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

$$H_a: \mu_1 < \mu_2 \ , or$$

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

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 of Difference

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

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

d.f.

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:

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

Prob Level

The probability level, also known as the *p*-value 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 = (0.050)

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

Alternative Hypothesis	N1	N2	μ1	μ2	σ1	σ2	Power (α = 0.05)	Power (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

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 =.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 Section

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 of Difference	Z-Statistic	Prob Level	Reject H0 at α = 0.050
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

Alternative Hypothesis

The (unreported) null hypothesis is

$$H_0: \mu_1 - \mu_2 = Hypothesized\ Difference = 0$$

and the alternative hypotheses,

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

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

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

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:

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

$$H_a: \mu_1 < \mu_2 \ , or$$

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

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 of Difference

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

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

Prob Level

The probability level, also known as the *p*-value 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 = (0.050)

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 Means

Equivalence Tests of Means (H0: μ1 - μ2 ≤ -0.3 or μ1 - μ2 ≥ 0.3 versus Ha: Equivalence)
Estimated Difference (Mean1 - Mean2) = 1.8188

Variance Assumption	H01: μ1 - μ2 ≤ -0.3 vs. Ha1: μ1 - μ2 > -0.3 Test Statistic	Prob Level	H02: μ1 - μ2 ≥ 0.3 vs. Ha2: μ1 - μ2 < 0.3 Test Statistic	Prob Level	Overall Prob Level	Reject Overall H0, Conclude Equivalence? (α = 0.050)
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 Prob Level

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

Reject Overall H0, Conclude Equivalence?

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

DF1	DF2	F-Value	Prob Level	Reject Hypothesis of Equal Variance? (α = 0.050)
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.

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:

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

Prob Level

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

Reject Hypothesis of Equal Variance?

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