

Six Sigma: Advanced Parametric and Nonparametric Hypothesis Tests

Short Examples Series
using
Risk Simulator



For more information please visit:
www.realoptionsvaluation.com
or contact us at:
admin@realoptionsvaluation.com

Six Sigma – One and Two Sample Hypothesis Tests using T-Tests, Z-Tests, F-Tests, ANOVA, and Nonparametric Tests (Friedman, Kruskal Wallis, Lilliefors, and Runs Test)

File Names: *Six Sigma – Hypothesis Testing (Advanced Techniques)*

Location: *Modeling Toolkit | Six Sigma | Hypothesis Testing (Advanced Techniques)*

Brief Description: *This chapter illustrates how to use Modeling Toolkit's Statistical Tools to run single and two variable hypothesis tests using parametric and nonparametric techniques*

Requirements: *Modeling Toolkit, Risk Simulator*

This chapter is a follow up from the last chapter on hypothesis testing. As a primer to the more advanced applications described here, please first review the materials in the last chapter. Here we continue the quest of quality control and sampling to test one and two different variables to see if they are statistically significantly similar or different from some hypothesized value. The methodologies introduced in this chapter are more advanced than those presented in the last chapter, but build on similar concepts. All the statistical tests shown in this chapter are available in the Statistical Tools section of Modeling Toolkit.

A hypothesis test is a statistical test of the characteristics a population by testing a small sample collected. In most cases, the population to be studied might be too large, difficult, or expensive to be completely sampled (e.g., all 100 million registered voters in the United States in a particular election). Hence, a smaller sample (e.g., a random sample of 1,100 voters from 20 cities) is collected, and the sample statistics are tabulated. Then, using hypothesis tests, the characteristics of the entire population can be inferred from this small sample. Modeling Toolkit allows the user to test one-variable, two-variable, and multiple-variable hypotheses tests using t-tests, Z-tests, and analysis of variance (ANOVA) techniques.

To perform a hypothesis test, first set up the null hypothesis (H_0) and the alternate hypothesis (H_a). Here are four quick rules:

- a) Always set up the alternate hypothesis first, then the null hypothesis
- b) The alternate hypothesis always has these signs: $>$ or $<$ or \neq
- c) The null hypothesis always has these signs: \geq or \leq or $=$
- d) If the alternate hypothesis is \neq then it's a 2-tailed test; if $<$ then it is a left (one) tailed test, if $>$ then it is a right (one) tailed test

If the p-value is less than the significance level (the significance level α is selected by the user and is usually 0.10, 0.05, or 0.01) tested, reject the null hypothesis and accept the alternate hypothesis.

Two-Tailed Hypothesis Test

A two-tailed hypothesis tests the null hypothesis of whether the population median of the sample dataset is statistically identical to the hypothesized median. The alternative hypothesis is that the real population median is statistically different from the hypothesized median when tested using the sample dataset. If the calculated p-value is less than or equal to the alpha significance value, then reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Right-Tailed Hypothesis Test

A right-tailed hypothesis tests the null hypothesis such that the population median of the sample dataset is statistically less than or equal to the hypothesized median. The alternative hypothesis is that the real population median is statistically greater than the hypothesized median when tested using the sample dataset. If the calculated p-value is less than or equal to the alpha significance value then reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Left-Tailed Hypothesis Test

A left-tailed hypothesis tests the null hypothesis such that the population median of the sample dataset is statistically greater than or equal to the hypothesized median. The alternative hypothesis is that the real population median is statistically less than the hypothesized median when tested using the sample dataset. If the calculated p-value is less than or equal to the alpha significance value then reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Example: A light bulb manufacturing company wants to test the hypothesis that its bulbs can last on average, 1000 burning hours. Hence, the company randomly selects 20 sample bulbs. The burning hours are shown in Figure 1. At a 5% significance level, what is the conclusion?

	A
1	1100
2	1095
3	997
4	1010
5	1001
6	1000
7	995
8	995
9	1005
10	1070
11	998
12	950
13	1020
14	1035
15	1050
16	999
17	1021
18	998
19	1010
20	1002

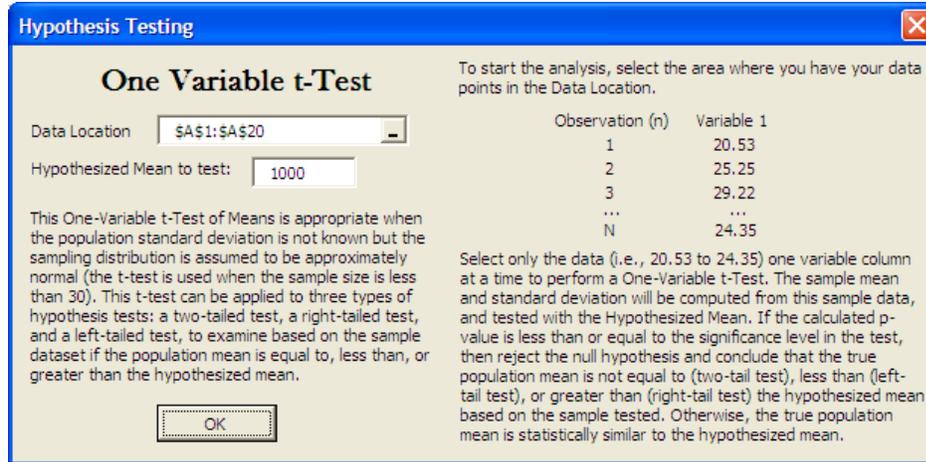


Figure 1: One-variable t-test

Figure 1 illustrates the sample data of 20 light bulbs randomly selected and tested using the One-Variable T-Test. Figure 2 illustrates the results obtained. The hypothesis tested is, of course, a one-tail right-tail test. As the sample size is less than 30, we use a one-variable t-test. Figure 1 also illustrates the data set up and requirements to run a one-variable t-test.

Clearly, the null hypothesis tested is such that the population average life of the light bulbs is $\leq 1,000$ hours, while the alternate hypothesis is such that the population life of the light bulbs $> 1,000$ hours. The calculated p-value for the right tail is 0.0220, which is less than the 5% significance (α) level. Hence we reject the null hypothesis and conclude that the population average of the life of the light bulbs exceed 1,000 hours.

Hypothesis Test (t-Test on the Population Mean of One Variable)

Statistical Summary

Statistics from Dataset:

Observations 20
 Sample Mean 1017.55
 Sample Standard Deviation 36.37

User Provided Statistics:

Hypothesized Mean 1000.00

Calculated Statistics:

t-Statistic 2.1582
 P-Value (right-tail) 0.0220
 P-Value (left-tailed) 0.9780
 P-Value (two-tailed) 0.0439

Null Hypothesis (Ho): $\mu = 1000.00$
 Alternate Hypothesis (Ha): $\mu < > 1000.00$

Notes: "<>" denotes "greater than" for right-tail, "less than" for left-tail, or "not equal to" for two-tail hypothesis tests.

Hypothesis Testing Summary

The one-variable t-test is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal (the t-test is used when the sample size is less than 30). This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. All three tests and their respective results are listed below for your reference.

Two-Tailed Hypothesis Test

A two-tailed hypothesis tests the null hypothesis H_0 such that the population mean is statistically identical to the hypothesized mean. The alternative hypothesis is that the real population mean is statistically different from the hypothesized mean when tested using the sample dataset. Using a t-test, the t value calculated is 2.1582 which corresponds to a p-value of 0.0439. This means that the hypothesized mean is significant at 10% and 5%. Therefore the null hypothesis is rejected, and the true population mean (tested through the sample mean 1017.55) is found to be statistically significantly different than the hypothesized mean of 1000.00.

Right-Tailed Hypothesis Test

A right-tailed hypothesis tests the null hypothesis H_0 such that the population mean is statistically less than or equal to the hypothesized mean. The alternative hypothesis is that the real population mean is statistically greater than the hypothesized mean when tested using the sample dataset. Using a t-test, the t value calculated is 2.1582 which corresponds to a p-value of 0.0220. This means that the hypothesized mean is significant at 10% and 5%. Therefore the null hypothesis is rejected, and that the true population mean (tested through the sample mean 1017.55) is found to be statistically significantly greater than the hypothesized mean of 1000.00.

Left-Tailed Hypothesis Test

A left-tailed hypothesis tests the null hypothesis H_0 such that the population mean is statistically greater than or equal to the hypothesized mean. The alternative hypothesis is that the real population mean is statistically less than the hypothesized mean when tested using the sample dataset. Using a t-test, the t value calculated is 2.1582 which corresponds to a p-value of 0.9780. This means that the hypothesized mean is not significant at any of the following significance levels: 1%, 5%, and 10%. Therefore the null hypothesis is not rejected, and that the true population mean (tested through the sample mean 1017.55) is found to be statistically not significantly less than the hypothesized mean of 1000.00.

Figure 2: Hypothesis test results

One-Variable Testing for Means (T-Test)

This One-Variable t-Test of Means is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal. The t-test is used when the sample size is less than 30. This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test, to examine based on the sample dataset if the population mean is equal to, less than, or greater than the hypothesized mean.

If the calculated p-value is less than or equal to the significance level in the test, then reject the null hypothesis and conclude that the true population mean is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized mean based on the sample tested. Otherwise, the true population mean is statistically similar to the hypothesized mean.

Data Requirements

For the **One-Variable T-Tests**, create data tables such as the one in Figure 3, and select the data area in blue **one variable** at a time (e.g., to test Variable One, select the data from A2 to A29, or data points 10 to 66). Repeat this process for other variables. To extend the data set, just add more observations (rows) or more variables (columns), as illustrated in Figures 3 and 4.

	A	B	C	D	E	F	G	H	I	J
1	One	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten
2	10.00	26.00	55.00	88.00	99.00	10.00	23.00	10.00	10.00	10.00
3	13.00	28.00	57.00	89.00	104.00	13.00	13.00	43.00	13.00	13.00
4	14.00	29.00	66.00	90.00	110.00	14.00	14.00	14.00	3.00	14.00
5	15.00	30.00	78.00	93.00	111.00	15.00	15.00	15.00	15.00	43.00
6	18.00	33.00	79.00	99.00	145.00	18.00	18.00	18.00	32.00	18.00
7	6.00	37.00	88.00	99.00	155.00	19.00	19.00	19.00	24.00	32.00
8	87.00	39.00	89.00	104.00	169.00	19.00	19.00	19.00	55.00	19.00
9	21.00	44.00	90.00	110.00	188.00	21.00	23.00	21.00	3.00	21.00
10	23.00	44.00	93.00	111.00	190.00	22.00	24.00	22.00	3.00	22.00
11	34.00	46.00	99.00	145.00	200.00	21.00	54.00	21.00	22.00	21.00
12	26.00	48.00	99.00	221.00	202.00	26.00	26.00	26.00	23.00	26.00
13	28.00	55.00	104.00	169.00	212.00	28.00	28.00	28.00	28.00	34.00
14	29.00	57.00	110.00	188.00	345.00	29.00	29.00	29.00	56.00	29.00
15	30.00	66.00	147.00	190.00	280.00	30.00	30.00	30.00	30.00	30.00
16	33.00	78.00	145.00	200.00	56.00	34.00	33.00	33.00	33.00	33.00
17	23.00	79.00	155.00	202.00	300.00	37.00	23.00	32.00	37.00	11.00
18	39.00	88.00	169.00	212.00	300.00	39.00	32.00	39.00	75.00	39.00
19	44.00	89.00	199.00	222.00	400.00	44.00	44.00	44.00	44.00	33.00
20	44.00	90.00	190.00	280.00	401.00	44.00	44.00	44.00	44.00	44.00
21	46.00	93.00	200.00	300.00	555.00	46.00	23.00	46.00	46.00	4.00
22	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	23.00
23	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
24	57.00	104.00	222.00	400.00	587.00	57.00	13.00	57.00	57.00	57.00
25	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00
26	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	48.00
27	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
28	57.00	104.00	222.00	400.00	587.00	57.00	57.00	57.00	57.00	57.00
29	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00

Figure 3: Data requirements for a one-variable t-test

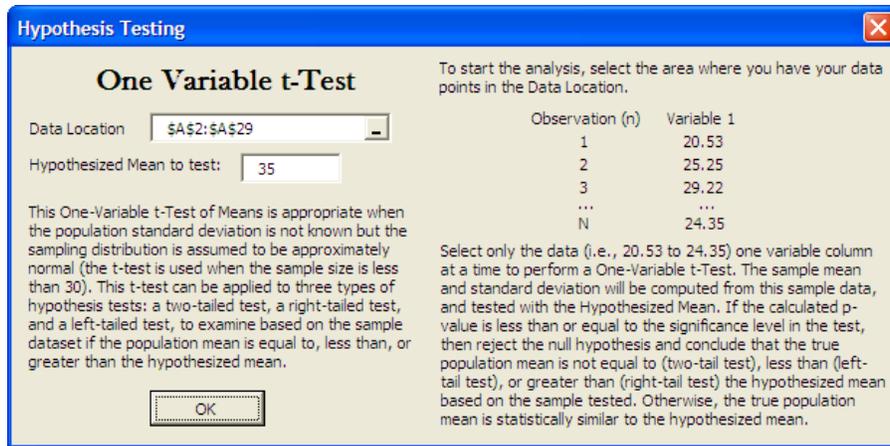


Figure 4: Running a one-variable t-test

One-Variable Testing for Means (Z-Test)

The one-variable Z-test is appropriate when the population standard deviation is known, and the sampling distribution is assumed to be approximately normal. This applies when the number of data points exceeds 30. This Z-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test, to examine based on the sample dataset if the population mean is equal to, less than, or greater than the hypothesized mean.

If the calculated p-value is less than or equal to the significance level in the test, then reject the null hypothesis and conclude that the true population mean is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized mean based on the sample tested. Otherwise, the true population mean is statistically similar to the hypothesized mean.

Data Requirements

For the **One-Variable Z-Test**, create data tables such as the one in Figure 5, and select the data area in blue **one variable** at a time (e.g., to test Variable Eleven, select the data from A2 to A34, or data points 10 to 58). Repeat this process for other variables. To extend the data set, just add more observations (rows) or more variables (columns) but make sure that you have at least 30 data points to run a Z-test (Figures 5 and 6).

	A	B	C	D	E
1	<i>Eleven</i>	<i>Twelve</i>	<i>Thirteen</i>	<i>Fourteen</i>	<i>Fifteen</i>
2	10.00	10.00	99.00	0.0360	0.0340
3	17.00	14.00	36.00	0.0990	0.0360
4	14.00	14.00	74.00	0.0360	0.0380
5	12.00	15.00	30.00	0.0740	0.0330
6	18.00	18.00	54.00	0.0300	0.0340
7	19.00	32.00	17.00	0.0540	0.0310
8	19.00	19.00	21.00	0.0170	0.0390
9	21.00	21.00	89.00	0.0320	0.0300
10	22.00	22.00	77.00	0.0890	0.0290
11	21.00	21.00	86.00	0.0770	0.0280
12	26.00	26.00	33.00	0.0860	0.0270
13	28.00	28.00	11.00	0.0330	0.0290
14	29.00	29.00	74.00	0.0900	0.0300
15	30.00	30.00	7.00	0.0740	0.0310
16	33.00	22.00	54.00	0.0070	0.0330
17	44.00	53.00	98.00	0.0540	0.0290
18	39.00	39.00	50.00	0.0980	0.0280
19	44.00	44.00	86.00	0.0500	0.0270
20	44.00	44.00	90.00	0.0860	0.0430
21	46.00	46.00	65.00	0.0900	0.0450
22	21.00	48.00	20.00	0.0650	0.0410
23	55.00	55.00	17.00	0.0200	0.0440
24	57.00	57.00	45.00	0.0170	0.0410
25	66.00	66.00	28.00	0.0450	0.0420
26	48.00	48.00	29.00	0.0890	0.0390
27	55.00	55.00	19.00	0.0770	0.0380
28	57.00	57.00	32.00	0.0450	0.0390
29	66.00	66.00	27.00	0.0210	0.0370
30	57.00	59.00	25.00	0.0450	0.0330
31	48.00	55.00	26.00	0.0330	0.0340
32	76.00	54.00	21.00	0.0230	0.0310
33	63.00	68.00	22.00	0.0200	0.0320
34	58.00	65.00	20.00	0.0210	0.0300

Figure 5: Data requirements for a one-variable z-test

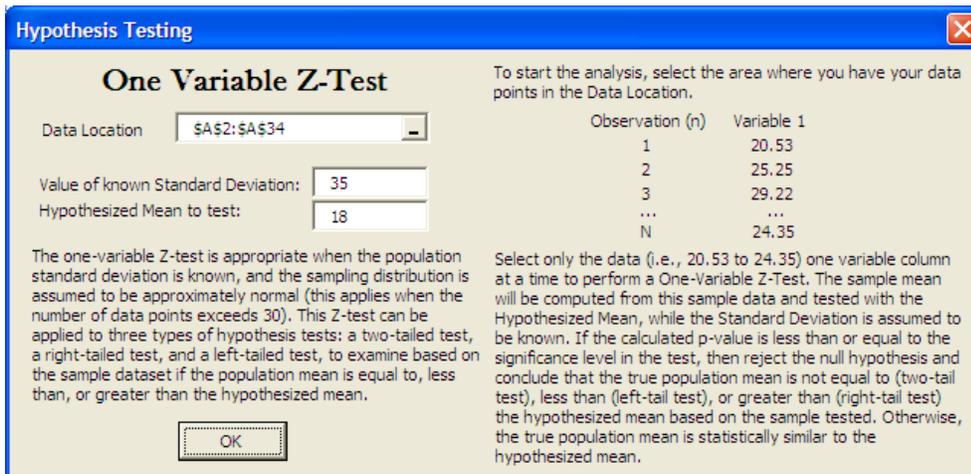


Figure 6: Running a one-variable z-test

One-Variable Testing for Proportions (Z-Test)

The one-variable Z-test for proportions is appropriate when the sampling distribution is assumed to be approximately normal. This applies when the number of data points exceeds 30, and when the number of data points, N , multiplied by the hypothesized population proportion mean, P , is greater than or equal to five, $NP \geq 5$. The data used in the analysis have to be proportions and be between 0 and 1. This Z-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test, to examine based on the sample dataset if the population mean is equal to, less than, or greater than the hypothesized mean.

If the calculated p-value is less than or equal to the significance level in the test, then reject the null hypothesis and conclude that the true population mean is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized mean based on the sample tested. Otherwise, the true population mean is statistically similar to the hypothesized mean.

For data requirements, see One Variable Testing for Means (Z-Test). However, make sure that the data are proportions between 0 and 1 (not including 1).

Two Variables with Dependent Means (T-Test)

The two-variable dependent t-test is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal. The t-test is used when the sample size is less than 30. In addition, this test is specifically formulated for testing the same or similar samples before and after an event (e.g., measurements taken before a medical treatment are compared against those measurements taken after the treatment to see if there is a difference). This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test.

Suppose that a new heart medication was administered to 100 patients ($N = 100$) and the heart rates before and after the medication was administered were measured. The Two Dependent Variables t-Test can be applied to test if the new medication is effective by testing to see if there is a statistically different "before and after" averages. The Dependent Variables test is used here because there is only a single sample collected (the heartbeats of the same patients were measured before and after administration of the new drug).

The two-tailed null hypothesis tests that the true population's mean of the difference between the two variables is zero, versus the alternate hypothesis that the difference is statistically different from zero. The right-tail null hypothesis test is such that the differences in the population means (first mean less second mean) is statistically less than or equal to zero (which is identical to saying that the mean of the first sample is less than or equal to the mean of the second sample). The alternative hypothesis is that the mean difference of the real populations is statistically greater than zero when tested using the sample dataset (which is identical to saying that the mean of the first sample is greater than the mean of the second sample). The left-tail null hypothesis test is such that the differences in the population means (first mean less second mean) is statistically greater than or equal to zero (which is the same as saying that the mean of the first sample is greater than or equal to the mean of the second sample). The alternative hypothesis is that the mean difference of the real population is statistically less than zero when tested using the sample dataset (which is identical to saying that the mean of the first sample is less than the mean of the second sample).

If the calculated p-value is less than or equal to the significance level in the test, reject the null hypothesis and conclude that the true population difference of the population means is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) zero based on the sample tested. Otherwise, the true population mean is statistically similar to the hypothesized mean.

Data Requirements

*For the **Two-Variable T-Tests**, create data tables such as the one shown in Figure 7, and select the data area in blue two variables at a time (e.g., to test Variables Eight and Nine together, select the data from H2 to I29 or the data points 10 to 66). Repeat this process for other pairs of variables. To extend the existing data set, just add more observations (rows) or more variables (columns). See Figures 7 and 8.*

	A	B	C	D	E	F	G	H	I	J
1	One	Two	Three	Four	Five	Six	Seven	Eight	Nine	Ten
2	10.00	26.00	55.00	88.00	99.00	10.00	23.00	10.00	10.00	10.00
3	13.00	28.00	57.00	89.00	104.00	13.00	13.00	43.00	13.00	13.00
4	14.00	29.00	66.00	90.00	110.00	14.00	14.00	14.00	3.00	14.00
5	15.00	30.00	78.00	93.00	111.00	15.00	15.00	15.00	15.00	43.00
6	18.00	33.00	79.00	99.00	145.00	18.00	18.00	18.00	32.00	18.00
7	6.00	37.00	88.00	99.00	155.00	19.00	19.00	19.00	24.00	32.00
8	87.00	39.00	89.00	104.00	169.00	19.00	19.00	19.00	55.00	19.00
9	21.00	44.00	90.00	110.00	188.00	21.00	23.00	21.00	3.00	21.00
10	23.00	44.00	93.00	111.00	190.00	22.00	24.00	22.00	3.00	22.00
11	34.00	46.00	99.00	145.00	200.00	21.00	54.00	21.00	22.00	21.00
12	26.00	48.00	99.00	221.00	202.00	26.00	26.00	26.00	23.00	26.00
13	28.00	55.00	104.00	169.00	212.00	28.00	28.00	28.00	28.00	34.00
14	29.00	57.00	110.00	188.00	345.00	29.00	29.00	29.00	56.00	29.00
15	30.00	66.00	147.00	190.00	280.00	30.00	30.00	30.00	30.00	30.00
16	33.00	78.00	145.00	200.00	56.00	34.00	33.00	33.00	33.00	33.00
17	23.00	79.00	155.00	202.00	300.00	37.00	23.00	32.00	37.00	11.00
18	39.00	88.00	169.00	212.00	300.00	39.00	32.00	39.00	75.00	39.00
19	44.00	89.00	199.00	222.00	400.00	44.00	44.00	44.00	44.00	33.00
20	44.00	90.00	190.00	280.00	401.00	44.00	44.00	44.00	44.00	44.00
21	46.00	93.00	200.00	300.00	555.00	46.00	23.00	46.00	46.00	4.00
22	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	23.00
23	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
24	57.00	104.00	222.00	400.00	587.00	57.00	13.00	57.00	57.00	57.00
25	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00
26	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	48.00
27	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
28	57.00	104.00	222.00	400.00	587.00	57.00	57.00	57.00	57.00	57.00
29	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00

Figure 7: Data requirements for a two-variable dependent means t-test

Hypothesis Testing
✕

Two Dependent Variables t-Test

Data Location:

The two-variable dependent t-test is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal (the t-test is used when the sample size is less than 30). In addition, this test is specifically formulated for testing the same or similar samples before and after an event (e.g., measurements taken before a medical treatment are compared against those measurements taken after the treatment to see if there is a difference). This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test.

The two-tailed null hypothesis tests that the true population's mean of the difference between the two variables is zero, versus the alternate hypothesis that the difference is statistically different from zero. The right-tail null hypothesis test is such that the differences in the population means (first mean less second mean) is statistically less than or equal to zero (which is identical to saying that mean of the first sample is less than or equal to the mean of the second sample). The alternative hypothesis is that the real populations' mean difference is statistically greater than zero when tested using the sample dataset (which is identical to saying that the mean of the first sample is greater than the mean of the second sample). The left-tail null hypothesis test is such that the differences in the population means (first mean less second mean) is statistically greater than or equal to zero (which is identical to saying that the mean of the first sample is greater than or equal to the mean of the second sample). The alternative hypothesis is that the real populations' mean difference is statistically less than zero when tested using the sample dataset (which is identical to saying that the mean of the first sample is less than the mean of the second sample).

To start the analysis, select the area where you have your data in the Data Location. The data for each Treatment should be arranged in different columns, for example:

Observation (n)	Variable 1	Variable 2
1	20.53	16.77
2	25.25	19.33
3	29.22	18.75
...
N	24.35	19.10

Select only the data area (e.g., from 20.53 to 19.10). In this illustration, suppose that a new heart medication was administered to 100 patients (N = 100) and the heart rates before and after the medication was administered were measured. The Two Dependent Variables t-Test can be applied to test if the new medication is effective by testing to see if there is a statistically different "before and after" averages. The Dependent Variables test is used here because there is only a single sample collected (the same patients' heartbeats were measured before and after the new drug administration).

If the calculated p-value is less than or equal to the significance level in the test, then reject the null hypothesis and conclude that the true population difference of the population means is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) zero based on the sample tested. Otherwise, the true difference of the population means is statistically similar to the hypothesized mean of zero.

Figure 8: Running a two-variable dependent means t-test

Two Variables that are Independent with Equal Variances (T-Test)

The two-variable t-test with equal variances is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal. The t-test is used when the sample size is less than 30. In addition, the two independent samples are assumed to have similar variances.

For illustration, suppose that a new engine design is tested against an existing engine design to see if there is a statistically significant difference between the two. The t-Test on Two (Independent) Variables with Equal Variances can be applied. This test is used here because there are two distinctly different samples collected here (new engine and existing engine) but the variances of both samples are assumed to be similar (the means may or may not be similar but the fluctuations around the mean are assumed to be similar).

This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the mean difference (HMD) of the population between the two variables is statistically identical to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean equals the second mean. The alternative hypothesis is that the difference between the real population means is statistically different from the hypothesized mean differences when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean does not equal the second mean.

A right-tailed hypothesis tests the null hypothesis H_0 such that the population mean difference between the two variables is statistically less than or equal to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean is less than or equals the second mean. The alternative hypothesis is that the real difference between population means is statistically greater than the hypothesized mean differences when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean is greater than the second mean.

A left-tailed hypothesis tests the null hypothesis H_0 such that the differences between the population means of the two variables is statistically greater than or equal to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean is greater than or equals the second mean. The alternative hypothesis is that the real difference between population means is statistically less

than the hypothesized mean difference when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean is less than the second mean.

If the calculated p-value is less than or equal to the significance level in the test, reject the null hypothesis and conclude that the true population difference of the population means is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) HMD based on the sample tested. Otherwise, the true difference of the population means is statistically similar to the HMD.

For data requirements, see Two Variables with Dependent Means (T-Test).

Two Variables that are Independent with Unequal Variances (T-Test)

The two-variable t-test with unequal variances (the population variance of sample 1 is expected to be different from the population variance of sample 2) is appropriate when the population standard deviation is not known but the sampling distribution is assumed to be approximately normal. The t-test is used when the sample size is less than 30. In addition, the two independent samples are assumed to have similar variances.

To illustrate, suppose that a new customer relationship management (CRM) process is being evaluated for its effectiveness. The customer satisfaction ranking between two hotels (one with and the other without CRM implemented) are collected. The t-Test on Two (Independent) Variables with Unequal Variances can be applied. This test is used here because there are two distinctly different samples collected (customer survey results of two different hotels) and the variances of both samples are assumed to be dissimilar (due to the difference in geographical location plus the different demographics and psychographics of customers at both properties).

This t-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the population mean differences between the two variables is statistically identical to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean equals the second mean. The alternative hypothesis is that the real difference between the population means is statistically different from the hypothesized mean differences when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean does not equal the second mean.

A right-tailed hypothesis tests the null hypothesis H_0 such that the difference between the means of the populations of the two variables is statistically less than or equal to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean is less than or equals the second mean. The alternative hypothesis is that the mean of the differences of the real populations is statistically greater than the hypothesized mean difference when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean is greater than the second mean.

A left-tailed hypothesis tests the null hypothesis H_0 such that the difference between the population means of the two variables is statistically greater than or equal to the hypothesized mean differences. If HMD is set to zero, this is the same as saying that the first mean is greater than or equals the second mean. The alternative hypothesis is that the real difference between population means is statistically less than the hypothesized mean difference when tested using the sample dataset. If HMD is set to zero, this is the same as saying that the first mean is less than the second mean.

If the calculated p-value is less than or equal to the significance level in the test, reject the null hypothesis and conclude that the true population difference of the population means is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized mean based on the sample tested. Otherwise, the true difference of the population means is statistically similar to the hypothesized mean.

For data requirements, see Two Variables with Dependent Means (T-Test).

Two Variables that are Independent testing for Means (Z-Test)

The two-variable Z-test is appropriate when the population standard deviations are known for the two samples, and the sampling distribution of each variable is assumed to be approximately normal. This applies when the number of data points of each variable exceeds 30.

To illustrate, suppose that a market research was conducted on two different markets and the sample collected is large (N must exceed 30 for both variables). The researcher is interested in testing whether there is a statistically significant different between the two markets. Further suppose that such a market survey has been performed many times in the past and the population standard deviations are known. A Two Independent Variables Z-Test can be applied because the sample size exceeds 30 on each market, and the population standard deviations are known.

This Z-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically identical to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically different from the hypothesized mean when tested using the sample dataset.

A right-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically less than or equal to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically greater than the hypothesized mean when tested using the sample dataset.

A left-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically greater than or equal to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically less than the hypothesized mean when tested using the sample dataset.

Data Requirements

*For the **Two Independent Variables Z-Test**, create data tables such as the one in Figure 9, and select the data area in blue **two variables** at a time (e.g., select the data from A2 to B34, or data points 10 to 65). To extend the data set, just add more observations (rows) or more variables (columns) but make sure that you have at least 30 data points in each variable to run a Z-test. See Figure 10.*

	A	B	C	D	E
1	<i>Eleven</i>	<i>Twelve</i>	<i>Thirteen</i>	<i>Fourteen</i>	<i>Fifteen</i>
2	10.00	10.00	99.00	0.0360	0.0340
3	17.00	14.00	36.00	0.0990	0.0360
4	14.00	14.00	74.00	0.0360	0.0380
5	12.00	15.00	30.00	0.0740	0.0330
6	18.00	18.00	54.00	0.0300	0.0340
7	19.00	32.00	17.00	0.0540	0.0310
8	19.00	19.00	21.00	0.0170	0.0390
9	21.00	21.00	89.00	0.0320	0.0300
10	22.00	22.00	77.00	0.0890	0.0290
11	21.00	21.00	86.00	0.0770	0.0280
12	26.00	26.00	33.00	0.0860	0.0270
13	28.00	28.00	11.00	0.0330	0.0290
14	29.00	29.00	74.00	0.0900	0.0300
15	30.00	30.00	7.00	0.0740	0.0310
16	33.00	22.00	54.00	0.0070	0.0330
17	44.00	53.00	98.00	0.0540	0.0290
18	39.00	39.00	50.00	0.0980	0.0280
19	44.00	44.00	86.00	0.0500	0.0270
20	44.00	44.00	90.00	0.0860	0.0430
21	46.00	46.00	65.00	0.0900	0.0450
22	21.00	48.00	20.00	0.0650	0.0410
23	55.00	55.00	17.00	0.0200	0.0440
24	57.00	57.00	45.00	0.0170	0.0410
25	66.00	66.00	28.00	0.0450	0.0420
26	48.00	48.00	29.00	0.0890	0.0390
27	55.00	55.00	19.00	0.0770	0.0380
28	57.00	57.00	32.00	0.0450	0.0390
29	66.00	66.00	27.00	0.0210	0.0370
30	57.00	59.00	25.00	0.0450	0.0330
31	48.00	55.00	26.00	0.0330	0.0340
32	76.00	54.00	21.00	0.0230	0.0310
33	63.00	68.00	22.00	0.0200	0.0320
34	58.00	65.00	20.00	0.0210	0.0300

Figure 9: Data requirements for a two-variable independent z-test

Hypothesis Testing ✖

Two Independent Variables Z-Test

Data Location:

Known Standard Deviation (Sample 1):

Known Standard Deviation (Sample 2):

Hypothesized Mean Difference:

The two-variable Z-test is appropriate when the population standard deviations are known for the two samples, and the sampling distribution of each variable is assumed to be approximately normal (this applies when the number of data points of each variable exceeds 30).

This Z-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically identical to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically different from the hypothesized mean when tested using the sample dataset.

A right-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically less than or equal to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically greater than the hypothesized mean when tested using the sample dataset.

A left-tailed hypothesis tests the null hypothesis H_0 such that the difference between the two population means is statistically greater than or equal to the hypothesized mean. The alternative hypothesis is that the real difference between the two population means is statistically less than the hypothesized mean when tested using the sample dataset.

To start the analysis, select the area where you have your data in the Data Location. The data for each variable should be arranged in different columns, for example:

Observation (n)	Variable 1	Variable 2
1	20.53	16.77
2	25.25	19.33
3	29.22	18.75
...
N	24.35	19.10

Select only the data area (e.g., from 20.53 to 19.10). In this illustration, suppose that a market research was conducted on two different markets and the sample collected is large (N must exceed 30 for both variables) and the researcher is interested in testing whether there is a statistically significant difference between the two markets. Further suppose that such a market survey has been performed many times in the past and the population standard deviations are known. A Two Independent Variables Z-Test can be applied because the sample size exceeds 30 on each market, and the population standard deviations are known.

If the calculated p-value is less than or equal to the significance level in the test, then reject the null hypothesis and conclude that the true population difference of the population means is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized mean based on the sample tested. Otherwise, the true difference of the population means is statistically similar to the hypothesized mean.

Figure 10: Running a two-variable independent z-test

Two Variables that are Independent testing for Proportions (Z-Test)

The two-variable Z-test on proportions is appropriate when the sampling distribution is assumed to be approximately normal. This applies when the number of data points of both samples exceeds 30. Further, the data should all be proportions and be between 0 and 1.

To illustrate, suppose that brand research was conducted on two different headache pills and the sample collected is large (N must exceed 30 for both variables). The researcher is interested in testing whether there is a statistically significant difference between the proportion of headache sufferers of both samples using the different headache medication. A Two Independent Variables Z-Test for Proportions can be applied because the sample size exceeds 30 on each market, and the data collected are proportions.

This Z-test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the difference in the population proportion is statistically identical to the hypothesized difference (if the hypothesized difference is set to zero, the null hypothesis tests if the population proportions of the two samples are identical). The alternative hypothesis is that the real difference in population proportions is statistically different from the hypothesized difference when tested using the sample dataset.

A right-tailed hypothesis tests the null hypothesis H_0 such that the difference in the population proportion is statistically less than or equal to the hypothesized difference. If the hypothesized difference is set to zero, the null hypothesis tests if population proportion of Sample 1 is equal or less than the population proportion of Sample 2. The alternative hypothesis is that the real difference in population proportions is statistically greater than the hypothesized difference when tested using the sample dataset.

A left-tailed hypothesis tests the null hypothesis H_0 such that the difference in the population proportion is statistically greater than or equal to the hypothesized difference. If the hypothesized difference is set to zero, the null hypothesis tests if population proportion of Sample 1 is equal or greater than the population proportion of Sample 2. The alternative hypothesis is that the real difference in population proportions is statistically less than the hypothesized difference when tested using the sample dataset.

Data Requirements

*For the **Two Independent Variables Z-Test for Proportions**, create data tables such as the one in Figure 11, and select the data area in blue **two variables** at a time (e.g., select the data from D2 to E34, or data*

points 0.0360 to 0.0300). To extend the data set, just add more observations (rows) or more variables (columns) but make sure that the data are proportions between 0 and 1 (not inclusive of 1). Finally, make sure that you have at least 30 data points to run a Z-test (see Figures 11 and 12).

	A	B	C	D	E
1	Eleven	Twelve	Thirteen	Fourteen	Fifteen
2	10.00	10.00	99.00	0.0360	0.0340
3	17.00	14.00	36.00	0.0990	0.0360
4	14.00	14.00	74.00	0.0360	0.0380
5	12.00	15.00	30.00	0.0740	0.0330
6	18.00	18.00	54.00	0.0300	0.0340
7	19.00	32.00	17.00	0.0540	0.0310
8	19.00	19.00	21.00	0.0170	0.0390
9	21.00	21.00	89.00	0.0320	0.0300
10	22.00	22.00	77.00	0.0890	0.0290
11	21.00	21.00	86.00	0.0770	0.0280
12	26.00	26.00	33.00	0.0860	0.0270
13	28.00	28.00	11.00	0.0330	0.0290
14	29.00	29.00	74.00	0.0900	0.0300
15	30.00	30.00	7.00	0.0740	0.0310
16	33.00	22.00	54.00	0.0070	0.0330
17	44.00	53.00	98.00	0.0540	0.0290
18	39.00	39.00	50.00	0.0980	0.0280
19	44.00	44.00	86.00	0.0500	0.0270
20	44.00	44.00	90.00	0.0860	0.0430
21	46.00	46.00	65.00	0.0900	0.0450
22	21.00	48.00	20.00	0.0650	0.0410
23	55.00	55.00	17.00	0.0200	0.0440
24	57.00	57.00	45.00	0.0170	0.0410
25	66.00	66.00	28.00	0.0450	0.0420
26	48.00	48.00	29.00	0.0890	0.0390
27	55.00	55.00	19.00	0.0770	0.0380
28	57.00	57.00	32.00	0.0450	0.0390
29	66.00	66.00	27.00	0.0210	0.0370
30	57.00	59.00	25.00	0.0450	0.0330
31	48.00	55.00	26.00	0.0330	0.0340
32	76.00	54.00	21.00	0.0230	0.0310
33	63.00	68.00	22.00	0.0200	0.0320
34	58.00	65.00	20.00	0.0210	0.0300

Figure 11: Data requirements for a two-variable z-test for proportions

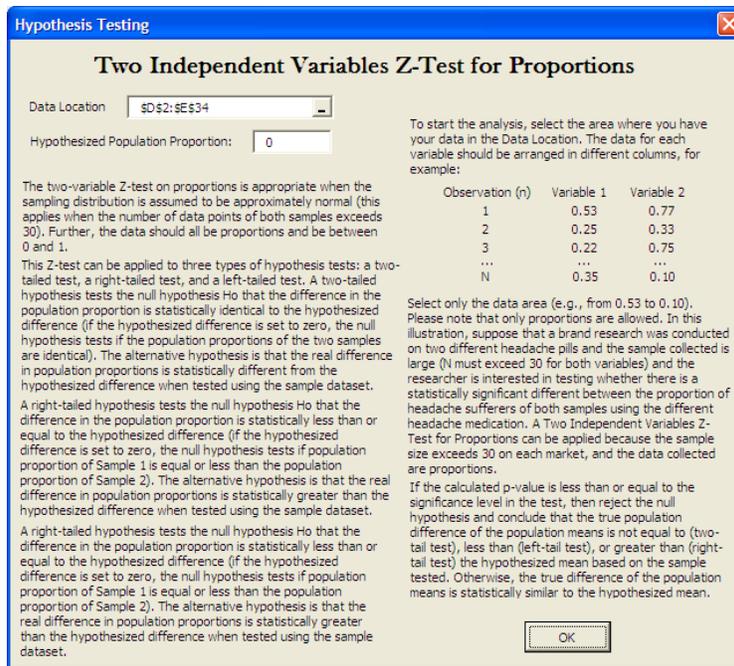


Figure 12: Running a two-variable z-test for proportions

Two Variables that are Independent testing for Variances (F-Test)

The two-variable F-test analyzes the variances from two samples. The population variance of sample 1 is tested with the population variance of sample 2 to see if they are equal. This test is appropriate when the population standard deviation is not known and the sampling distribution is assumed to be approximately normal.

The measurement of variation is a key issue in Six Sigma and quality control applications. In this illustration, suppose that the variation or variance around the units produced in a manufacturing process is compared to another process to determine which process is more variable and hence, less predictable in quality.

This F-test typically can be applied to a single hypothesis test: a two-tailed test. A two-tailed hypothesis tests the null hypothesis H_0 such that the population variance of the two variables is statistically identical. The alternative hypothesis is that the population variances are statistically different from one another when tested using the sample dataset.

If the calculated p-value is less than or equal to the significance level in the test, reject the null hypothesis and conclude that the true population variances of the two variables are not statistically equal to one another. Otherwise, the true population variances are statistically similar to each other.

For data requirements, see Two Variables that are Independent testing for Proportions (Z-Test).

The Basics of Nonparametric Methodologies

Nonparametric techniques make no assumptions about the specific shape or distribution from which the sample is drawn. This is different from the other hypotheses tests, such as ANOVA or t-Tests (parametric tests) where the sample is assumed to be drawn from a population that is normally or approximately normally distributed. If normality is assumed, the power of the test is higher due this normality restriction. However, if flexibility on distributional requirements is needed, nonparametric techniques are superior. In general, nonparametric methodologies provide these advantages over other parametric tests:

- Normality or approximate normality does not have to be assumed
- Fewer assumptions about the population are required, that is, nonparametric tests do not require that the population assume any specific distribution
- Smaller sample sizes can be analyzed
- Samples with nominal and ordinal scales of measurement can be tested
- Sample variances do not have to be equal, which is required in parametric tests

However, several caveats are worthy of mention:

- Compared to parametric tests, nonparametric tests use data less efficiently
- The power of nonparametric tests is lower than that of parametric tests

Therefore, if all the required assumptions are satisfied, it is better to use parametric tests. However, in reality, it may be difficult to justify these distributional assumptions, or small sample sizes may exist, requiring the need for nonparametric tests. Thus, nonparametric tests should be used when the data are nominal or ordinal, or when the data are interval or ratio but the normality assumption is not met.

The following lists each of the nonparametric tests available for use in the software.

Chi-Square Goodness of Fit Test

The Chi-Square test for goodness of fit is used to examine if a sample data set could have been drawn from a population having a specified probability distribution. The probability distribution tested here is the normal distribution. The null hypothesis tested such that the sample is randomly drawn from the normal distribution, versus the alternate hypothesis that the sample is not from a normal distribution. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Data Requirements

For the **Chi-Square Goodness of Fit Test**, create data tables such as the one shown in Figure 13, and select the data area in blue area (e.g., select the data from D6 to E13, or data points 800 to 4). To extend the data set, just add more observations (rows) as seen in Figures 13 and 14.

	C	D	E
4	Chi-Square Goodness of Fit		
5	Category	Upper Limit	Frequency
6	700-800	800	36
7	800-900	900	96
8	900-1000	1000	78
9	1000-1100	1100	48
10	1100-1200	1200	25
11	1200-1300	1300	10
12	1300-1400	1400	3
13	1400-1500	1500	4
14	<i>*For the Chi-Square Goodness of Fit Test, create</i>		
15	<i>data tables such as these, and select the data area</i>		
16	<i>in blue area (e.g., select the data from D6 to E13, or</i>		
17	<i>data points 800 to 4). To extend the data set, just</i>		
18	<i>add more observations (rows).</i>		

Figure 13: Data requirements for the chi-square goodness-of-fit test

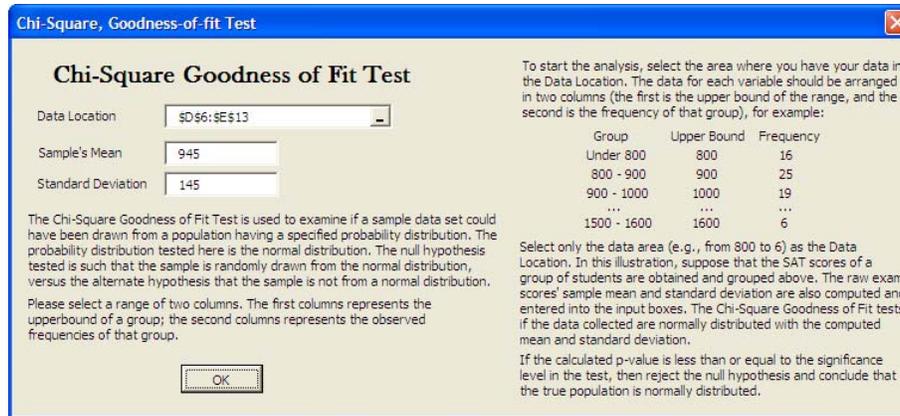


Figure 14: Running a chi-square goodness of fit test

Chi-Square Test of Independence

The Chi-Square test for independence examines two variables to see if there is some statistical relationship between them. This test is not used to find the exact nature of the relationship between the two variables but simply to test if the variables could be independent of each other. The null hypothesis tested is such that the variables are independent of each other, versus the alternate hypothesis that the variables are not independent of each other. The Chi-Square test looks at a table of observed frequencies and a table of expected frequencies. The amount of disparity between these two tables is calculated and compared with the Chi-Square test statistic. The observed frequencies reflect the cross-classification for members of a single sample, and the table of expected frequencies is constructed under the assumption that the null hypothesis is true.

Data Requirements

For the **Chi-Square Test of Independence**, create data tables such as the one in figure 15.15, and select the data area in blue area two variables at a time (e.g., select the data from A2 to B29, or data points 10 to 110). To extend the data set, just add more observations (rows) as seen in Figure 15.

	A	B
1	One	Two
2	10.00	26.00
3	13.00	28.00
4	14.00	29.00
5	15.00	30.00
6	18.00	33.00
7	6.00	37.00
8	87.00	39.00
9	21.00	44.00
10	23.00	44.00
11	34.00	46.00
12	26.00	48.00
13	28.00	55.00
14	29.00	57.00
15	30.00	66.00
16	33.00	78.00
17	23.00	79.00
18	39.00	88.00
19	44.00	89.00
20	44.00	90.00
21	46.00	93.00
22	48.00	99.00
23	55.00	99.00
24	57.00	104.00
25	66.00	110.00
26	48.00	99.00
27	55.00	99.00
28	57.00	104.00
29	66.00	110.00

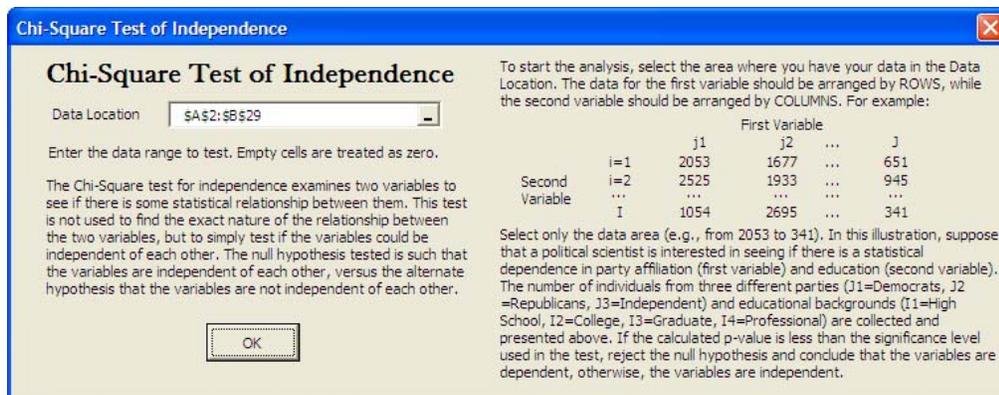


Figure 15: Data requirements and running the chi-square test of independence

Chi-Square Population Variance Test

The Chi-Square test for population variance is used for hypothesis testing and confidence interval estimation for a population variance. The population variance of a sample is typically unknown; hence, the need for quantifying this confidence interval. The population is assumed to be normally distributed.

Friedman's Test

The Friedman test is a form of nonparametric test, that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This method is the extension of the Wilcoxon Signed Rank test for paired samples. The corresponding parametric test is the Randomized Block Multiple Treatment ANOVA, but unlike the ANOVA, the Friedman test does not require that the dataset be randomly sampled from normally distributed populations with equal variances.

The Friedman test uses a two-tailed hypothesis test where the null hypothesis is such that the population medians of each treatment are statistically identical to the rest of the group. That is, there is no effect on the treatment among the different groups. The alternative hypothesis is such that the real population medians are statistically different from one another when tested using the sample dataset. That is, the

medians are statistically different, which means that there is a statistically significant effect among the different treatment groups. If the calculated p-value is less than or equal to the alpha significance value then reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Data Requirements

For the **Friedman's Test**, create data tables such as the one shown in Figure 16, and select the data area in blue area (e.g., select the data from C22 to F32, or data points Treatment 1 to 80). If selecting the headers, remember to select "Treat first row as headers." To extend the data set, just add more observations (rows). See Figure 16.

	B	C	D	E	F
22	<i>Blocks</i>	<i>Treatment 1</i>	<i>Treatment 2</i>	<i>Treatment 3</i>	<i>Treatment 4</i>
23	1	90	87	93	85
24	2	86	79	87	83
25	3	76	77	91	85
26	4	75	78	92	83
27	5	79	79	89	82
28	6	68	75	88	83
29	7	69	74	87	84
30	8	68	76	82	81
31	9	59	72	91	81
32	10	62	71	90	80

33
 34 **For the Friedman's Test, create data tables such as these, and select the data*
 35 *area in blue area (e.g., select the data from C22 to F32, or data points Treatment 1*
 36 *to 80). If selecting the headers, remember to select "Treat first row as headers." To*
 37 *extend the data set, just add more observations (rows).*

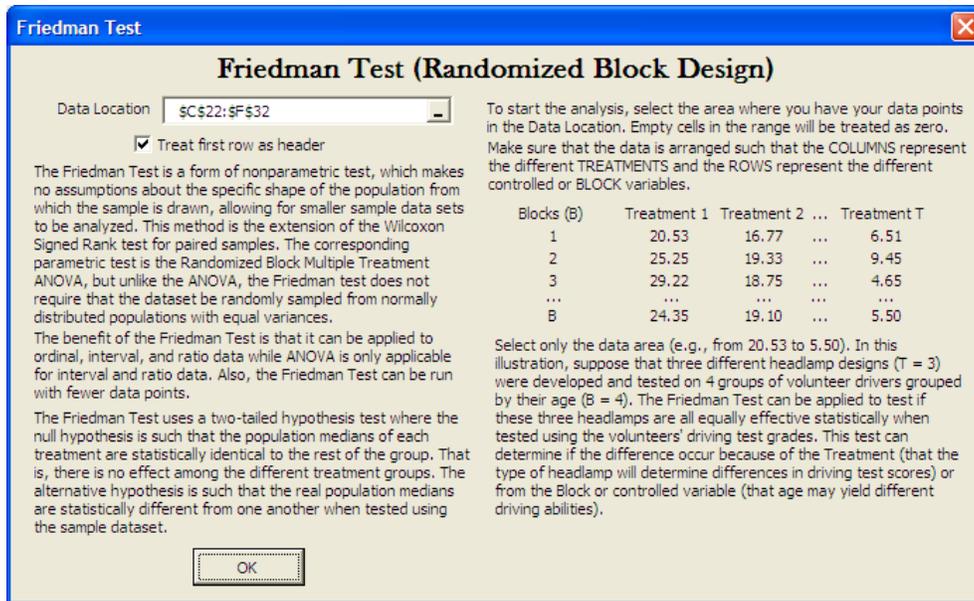


Figure 16: Data requirements and running the Friedman test for randomized block design

Kruskal-Wallis Test

The Kruskal-Wallis test is a form of nonparametric test that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This method is the extension of the Wilcoxon Signed Rank test for paired samples that compares the statistical properties of more than two independent samples. The corresponding parametric test is the One-Way ANOVA, but unlike the ANOVA, the Kruskal-Wallis test does not require that the dataset be randomly sampled from normally distributed populations with equal variances. The Kruskal-Wallis test is a two-tailed hypothesis test where the null hypothesis is such that the population medians of each treatment are statistically identical to the rest of the group. That is, there are no different effects among the different treatment groups. The alternative hypothesis is such that the real population medians are statistically different from one another when tested using the sample dataset. That is, the medians are statistically different, which means that there is a statistically significant effect among the different treatment groups. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

The benefit of the Kruskal-Wallis Test is that it can be applied to ordinal, interval, and ratio data; ANOVA, however, is applicable only for interval and ratio data. Also, the Kruskal-Wallis test can be run with fewer data points.

To illustrate, suppose that three different drug indications ($T = 3$) were developed and tested on 100 patients each ($N = 100$). The Kruskal-Wallis Test can be applied to test if these three drugs are all equally effective statistically. If the calculated p-value is less than or equal to the significance level used in the test, reject the null hypothesis and conclude that there is a significant difference among the different Treatments. Otherwise, the Treatments are all equally effective.

Data Requirements

For the **Kruskal-Wallis Test**, create data tables such as the one in Figure 17, and select the data area in blue area (e.g., select the data from C40 to F50, or data points Treatment 1 to 80). To extend the data set, just add more observations (rows) or more treatment variables to compare (columns). If selecting the variable names, remember to select “Treat first row as headers.”

	B	C	D	E	F
39	One-Way ANOVA and Kruskal-Wallis Test				
40	<i>Observations</i>	<i>Treatment 1</i>	<i>Treatment 2</i>	<i>Treatment 3</i>	<i>Treatment 4</i>
41	1	90	87	93	85
42	2	86	79	87	83
43	3	76	77	91	85
44	4	75	78	92	83
45	5	79	79	89	82
46	6	68	75	88	83
47	7	69	74	87	84
48	8	68	76	82	81
49	9	59	72	91	81
50	10	62	71	90	80

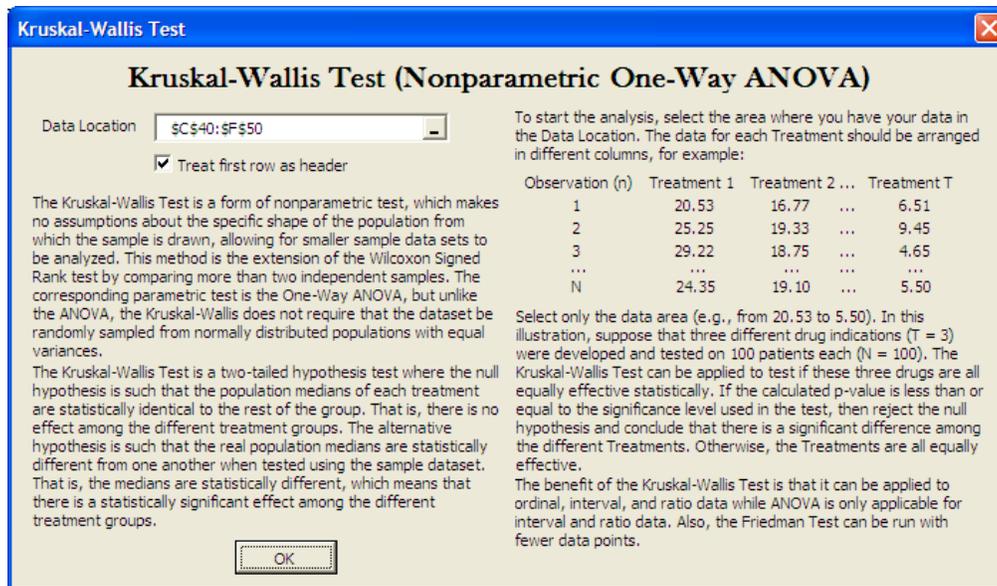


Figure 17: Data requirements and running the Kruskal-Wallis test for a nonparametric one-way ANOVA equivalence

Lilliefors Test

The Lilliefors test is a form of nonparametric test, that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This test evaluates the null hypothesis of whether the data sample was drawn from a normally distributed population, versus an alternate hypothesis that the data sample is not normally distributed. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis. This test relies on two cumulative frequencies: one derived from the sample data set, the second derived from a theoretical distribution based on the mean and standard deviation of the sample data. An alternative to this test is the Chi-Square test for normality, which requires more data points to run than the Lilliefors test.

Data Requirements

For the ***Lilliefors Test for Normality***, create data tables such as the one shown in Figure 18, and select the data area in blue area ***one variable*** at a time (e.g., select the data from A2 to A29, or data points 10 to 66). To extend the data set, just add more observations (rows).

	A
1	One
2	10.00
3	13.00
4	14.00
5	15.00
6	18.00
7	6.00
8	87.00
9	21.00
10	23.00
11	34.00
12	26.00
13	28.00
14	29.00
15	30.00
16	33.00
17	23.00
18	39.00
19	44.00
20	44.00
21	46.00
22	48.00
23	55.00
24	57.00
25	66.00
26	48.00
27	55.00
28	57.00
29	66.00

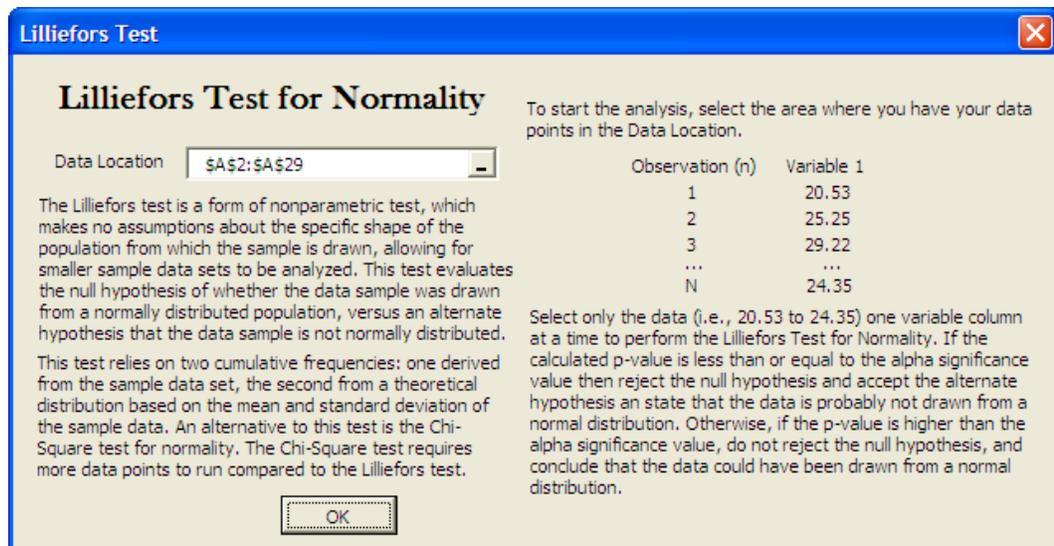


Figure 18: Data requirements and running the nonparametric Lilliefors test for normality

Runs Test

The Runs test is a form of nonparametric test that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This test evaluates the randomness of a series of observations by analyzing the number of runs it contains. A run is a consecutive appearance of one or more observations that are similar. The null hypothesis tested is whether the data sequence is random, versus the alternate hypothesis that the data sequence is not random. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Data Requirements

For the ***Runs Test for Randomness***, create data tables such as the one shown in Figure 19, and select the data area in blue area **multiple variables** at a time (e.g., select the data from A2 to J29, or data points 10 to 66). To extend the data set, just add more observations (rows) and variables (columns).

	A	B	C	D	E	F	G	H	I	J
1	<i>One</i>	<i>Two</i>	<i>Three</i>	<i>Four</i>	<i>Five</i>	<i>Six</i>	<i>Seven</i>	<i>Eight</i>	<i>Nine</i>	<i>Ten</i>
2	10.00	26.00	55.00	88.00	99.00	10.00	23.00	10.00	10.00	10.00
3	13.00	28.00	57.00	89.00	104.00	13.00	13.00	43.00	13.00	13.00
4	14.00	29.00	66.00	90.00	110.00	14.00	14.00	14.00	3.00	14.00
5	15.00	30.00	78.00	93.00	111.00	15.00	15.00	15.00	15.00	43.00
6	18.00	33.00	79.00	99.00	145.00	18.00	18.00	18.00	32.00	18.00
7	6.00	37.00	88.00	99.00	155.00	19.00	19.00	19.00	24.00	32.00
8	87.00	39.00	89.00	104.00	169.00	19.00	19.00	19.00	55.00	19.00
9	21.00	44.00	90.00	110.00	188.00	21.00	23.00	21.00	3.00	21.00
10	23.00	44.00	93.00	111.00	190.00	22.00	24.00	22.00	3.00	22.00
11	34.00	46.00	99.00	145.00	200.00	21.00	54.00	21.00	22.00	21.00
12	26.00	48.00	99.00	221.00	202.00	26.00	26.00	26.00	23.00	26.00
13	28.00	55.00	104.00	169.00	212.00	28.00	28.00	28.00	28.00	34.00
14	29.00	57.00	110.00	188.00	345.00	29.00	29.00	29.00	56.00	29.00
15	30.00	66.00	147.00	190.00	280.00	30.00	30.00	30.00	30.00	30.00
16	33.00	78.00	145.00	200.00	56.00	34.00	33.00	33.00	33.00	33.00
17	23.00	79.00	155.00	202.00	300.00	37.00	23.00	32.00	37.00	11.00
18	39.00	88.00	169.00	212.00	300.00	39.00	32.00	39.00	75.00	39.00
19	44.00	89.00	199.00	222.00	400.00	44.00	44.00	44.00	44.00	33.00
20	44.00	90.00	190.00	280.00	401.00	44.00	44.00	44.00	44.00	44.00
21	46.00	93.00	200.00	300.00	555.00	46.00	23.00	46.00	46.00	4.00
22	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	23.00
23	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
24	57.00	104.00	222.00	400.00	587.00	57.00	13.00	57.00	57.00	57.00
25	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00
26	48.00	99.00	202.00	300.00	565.00	48.00	48.00	48.00	48.00	48.00
27	55.00	99.00	212.00	300.00	577.00	55.00	55.00	55.00	55.00	55.00
28	57.00	104.00	222.00	400.00	587.00	57.00	57.00	57.00	57.00	57.00
29	66.00	110.00	280.00	401.00	599.00	66.00	66.00	66.00	66.00	66.00

Figure 19: Data requirements for the nonparametric Runs test for randomness

Wilcoxon Signed-Rank Test (One Variable)

The single variable Wilcoxon Signed Rank test is a form of nonparametric test that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This method looks at whether a sample dataset could have been randomly drawn from a particular population whose median is being hypothesized. The corresponding parametric test is the one-sample t-test, which should be used if the underlying population is assumed to be normal, providing a higher statistical power on the test. The Wilcoxon Signed Rank test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test. If the calculated Wilcoxon statistic is outside the critical limits for the specific significance level in the test, reject the null hypothesis and conclude that the true population median is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized median based on the sample tested. Otherwise, the true population median is statistically similar to the hypothesized median.

For data requirements, see One-Variable Testing for Means (T-Test).

Wilcoxon Signed-Rank Test (Two Variables)

The Wilcoxon Signed Rank test for paired variables is a form of nonparametric test that makes no assumptions about the specific shape of the population from which the sample is drawn, allowing for smaller sample data sets to be analyzed. This method looks at whether the median of the differences between the two paired variables is equal. This test is specifically formulated for testing the same or similar samples before and after an event (e.g., measurements taken before a medical treatment are compared against those measurements taken after the treatment to see if there is a difference). The corresponding parametric test is the two-sample t-test with dependent means, which should be used if the underlying population is assumed to be normal, providing a higher statistical power on the test. The Wilcoxon Signed Rank test can be applied to three types of hypothesis tests: a two-tailed test, a right-tailed test, and a left-tailed test.

To illustrate, suppose that a new engine design is tested against an existing engine design to see if there is a statistically significant difference between the two. The Paired Variable Wilcoxon Signed-Rank Test can be applied. If the calculated Wilcoxon statistic is outside the critical limits for the specific significance level in the test, reject the null hypothesis and conclude that the difference between the true population medians is not equal to (two-tail test), less than (left-tail test), or greater than (right-tail test) the hypothesized median difference based on the sample tested. Otherwise, the true population median is statistically similar to the hypothesized median.

For data requirements, see Two Dependent Variables Testing for Means (Z-Test).

Single-Factor Multiple Treatments ANOVA

The One-Way ANOVA for Single Factor with Multiple Treatments is an extension of the two-variable t-test, looking at multiple variables simultaneously. The ANOVA is appropriate when the sampling distribution is assumed to be approximately normal. ANOVA can be applied only to the two-tailed hypothesis test. A two-tailed hypothesis tests the null hypothesis such that the population means of each treatment are statistically identical to the rest of the group, which means that there is no effect among the different treatment groups. The alternative hypothesis is such that the real population means are statistically different from one another when tested using the sample dataset.

To illustrate, suppose that three different drug indications ($T = 3$) were developed and tested on 100 patients each ($N = 100$). The One-Way ANOVA can be applied to test if these three drugs are all equally

effective statistically. If the calculated p-value is less than or equal to the significance level used in the test, reject the null hypothesis and conclude that there is a significant difference among the different Treatments. Otherwise, the Treatments are all equally effective.

Data Requirements

For the **One-Way ANOVA** module, create tables such as the one in Figure 20, and select the data area in blue (15 to 25). You can extend the data by adding rows of observations and columns of treatments.

	B	C	D	E
4	One-Way ANOVA			
5	<i>Observations</i>	<i>Treatment 1</i>	<i>Treatment 2</i>	<i>Treatment 3</i>
6	1	15	10	18
7	2	20	15	19
8	3	19	11	23
9	4	14	14	25

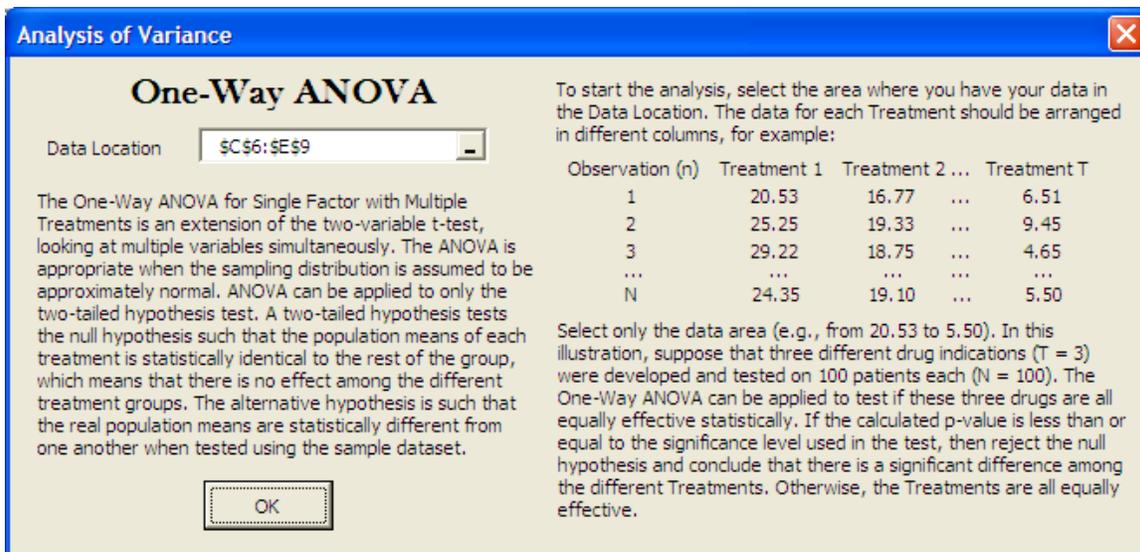


Figure 20: Data requirements and running a one-way ANOVA

Randomized Block Multiple Treatments ANOVA

The One-Way Randomized Block ANOVA is appropriate when the sampling distribution is assumed to be approximately normal and when there exists a Block variable for which ANOVA will Control (Block the effects of this variable by controlling it in the experiment). ANOVA can be applied only to the two-tailed hypothesis test. This analysis can test for the effects of both the Treatments as well as the

effectiveness of the Control or Block variable. If the calculated p-value for the Treatment is less than or equal to the significance level used in the test, reject the null hypothesis and conclude that there is a significant difference among the different Treatments. If the calculated p-value for the Block variable is less than or equal to the significance level used in the test, reject the null hypothesis and conclude that there is a significant difference among the different Block variables. To illustrate, suppose that three different headlamp designs ($T = 3$) were developed and tested on 4 groups of volunteer drivers grouped by their age ($B = 4$). The One-Way Randomized Block ANOVA can be applied to test if these three headlamps are all equally effective statistically when tested using the volunteers' driving test grades. Otherwise, the Treatments are all equally effective. This test can determine if the difference occur because of the Treatment (that the type of headlamp will determine differences in driving test scores) or from the Block or controlled variable (that age may yield different driving abilities).

Data Requirements

For the **Randomized Block ANOVA** module, create tables such as the one in Figure 21, and select the data area (e.g., values 90 to 73 in Figure 21). You can extend the data by adding rows of blocks and columns of treatments.

	B	C	D	E	F
17	Randomized Block Design ANOVA				
18	<i>Blocks</i>	<i>Treatment 1</i>	<i>Treatment 2</i>	<i>Treatment 3</i>	<i>Treatment 4</i>
19	1	90	87	93	85
20	2	86	79	87	83
21	3	76	74	77	73

Analysis of Variance
✖

One-Way Randomized
Block ANOVA

Data Location:

The One-Way Randomized Block ANOVA is appropriate when the sampling distribution is assumed to be approximately normal and when there exists a Block variable for which ANOVA will Control (Block the effects of this variable by controlling it in the experiment). ANOVA can be applied to only the two-tailed hypothesis test. This analysis can test for the effects of both the Treatments as well as the effectiveness of the Control or Block variable.

If the calculated p-value for the Treatment is less than or equal to the significance level used in the test, then reject the null hypothesis and conclude that there is a significant difference among the different Treatments. If the calculated p-value for the Block variable is less than or equal to the significance level used in the test, then reject the null hypothesis and conclude that there is a significant difference among the different Block variables.

To start the analysis, select the area where you have your data points in the Data Location.

Make sure that the data is arranged such that the COLUMNS represent the different TREATMENTS and the ROWS represent the different controlled or BLOCK variables.

Blocks (B)	Treatment 1	Treatment 2	...	Treatment T
1	20.53	16.77	...	6.51
2	25.25	19.33	...	9.45
3	29.22	18.75	...	4.65
...
B	24.35	19.10	...	5.50

Select only the data area (e.g., from 20.53 to 5.50). In this illustration, suppose that three different headlamp designs ($T = 3$) were developed and tested on 4 groups of volunteer drivers grouped by their age ($B = 4$). The One-Way Randomized Block ANOVA can be applied to test if these three headlamps are all equally effective statistically when tested using the volunteers' driving test grades. Otherwise, the Treatments are all equally effective. This test can determine if the difference occur because of the Treatment (that the type of headlamp will determine differences in driving test scores) or from the Block or controlled variable (that age may yield different driving abilities).

Figure 21: Data requirements and running a randomized block ANOVA

Two-Way ANOVA

The Two-Way ANOVA is an extension of the Single Factor and Randomized Block ANOVA by simultaneously examining the effects of two factors on the dependent variable, along with the effects of interactions between the different levels of these two factors. Unlike the randomized block design, this model examines the interactions between different levels of the factors, or independent variables. In a two-factor experiment, interaction exists when the effect of a level for one factor depends on which level of the other factor is present.

There are three sets of null and alternate hypotheses to be tested in the two-way analysis of variance.

The first test is on the first independent variable, where the null hypothesis is that no level of the first factor has an effect on the dependent variable. The alternate hypothesis is that there is at least one level of the first factor having an effect on the dependent variable. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

The second test is on the second independent variable, where the null hypothesis is that no level of the second factor has an effect on the dependent variable. The alternate hypothesis is that there is at least one level of the second factor having an effect on the dependent variable. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

The third test is on the interaction of both the first and second independent variables, where the null hypothesis is that there are no interacting effects between levels of the first and second factors. The alternate hypothesis is that there is at least one combination of levels of the first and second factors having an effect on the dependent variable. If the calculated p-value is less than or equal to the alpha significance value, reject the null hypothesis and accept the alternate hypothesis. Otherwise, if the p-value is higher than the alpha significance value, do not reject the null hypothesis.

Data Requirements

For the Two-Way ANOVA module, create tables such as the one in Figure 151.22, and select the data area (e.g., the values 804 to 835 in Figure 22). You can extend the data by adding rows of factors and columns of treatments. Note that the number of replications in the table above is 2 (i.e., two rows of

observations per Factor A type). Of course you can increase the number of replications as required. The number of replications has to be consistent if you wish to extend the data set.

	B	C	D	E	F
26	Two-Way ANOVA				
27			Factor B		
28					
29			<i>j = 1</i>	<i>j = 2</i>	<i>j = 3</i>
30					
31	Factor A	<i>i = 1</i>	804	836	804
32			816	828	808
33		<i>i = 2</i>	819	844	807
34			813	836	819
35		<i>i = 3</i>	820	814	819
36			821	811	829
37		<i>i = 4</i>	806	811	827
38		805	806	835	

Analysis of Variance ✖

Two-Way ANOVA

Data Location:

Number of replications:

The Two-Way ANOVA is an extension of the One-Way ANOVA and One-Way Randomized Block ANOVA by simultaneously examining the effects of Two Factors on the dependent variable, along with the effects of interactions between the different levels of these Two Factors. Unlike the Randomized Block design, this model examines the interactions between different levels of the factors, or independent variables. In a two-factor experiment, interaction exists when the effect of a level for one factor depends on which level of the other factor is present.

The hypothesis can be tested on Factor A, Factor B, and the interaction effects of A and B. Three p-values are calculated. If the calculated p-value is less than or equal to the significance level used in the test, then reject the null hypothesis and conclude that there is a significant difference among the different factors. Otherwise, the factors are all equally effective.

To start the analysis, select the area where you have your data in the Data Location. The data for the first Factor A should be arranged by ROWS, while the second Factor B should be arranged by COLUMNS. The Replication of each factor are arranged as sub-ROWS between factors. For example:

		Column Factor B			
		j1	j2	...	J
Row Factor A	i=1, r=1	20.53	16.77	...	6.51
	i=1, r=2	25.25	19.33	...	9.45
	i=2, r=1	23.61	14.75	...	8.55
	i=2, r=2	29.95	16.88	...	7.15

	I, r=1	24.35	19.10	...	5.50
I, r=2	22.15	18.00	...	7.64	

Select only the data area (e.g., from 20.53 to 7.64). In this illustration, suppose that a manufacturer is testing the tensile strength of three types of metals (J = 3) where each metal has five different thickness levels (I = 5). For each combination of metal-thickness, two samples are replicated (r = 2).

Figure 22: Data requirements and running a two-way ANOVA