Hypothesis Testing T-Test

Updated: July 1, 2015

Compares means from two data sets using a t-test

Category: Statistical Functions (https://msdn.microsoft.com/en-us/library/azure/dn905867.aspx)

Module Overview

You can use the **Hypothesis Testing T-Test** module to generate scores for three types of t-tests:

- Single sample t-test
- Paired t-test
- Unpaired t-test

In general, a t-test helps you compare whether two groups have different means. For example, suppose you are evaluating trial data for patients who received Drug A vs. patients who received Drug B, and you need to compare a recovery rate metric for both groups. The *null hypothesis* would assume that the recovery rate is the same in both groups, and furthermore, that the values for the recovery rate have a normal distribution in both two groups.

By using **Hypothesis Testing T-Test** and providing the columns that contain the recovery rates as input, you can get scores that indicate whether the difference is meaningful, which would signify that the null hypothesis should be rejected. The test takes into account factors such as how big the difference is between the values, the size of the sample (larger is better), and how big the standard deviation is (lower is better).

Hypothesis Testing T-Test indicates whether the null hypothesis is TRUE or FALSE in the right-hand output, and provides the confidence (*P*) score from the t-test in the left-hand output.

How to Configure Hypothesis T-Test

1. Decide which kind of t-test to perform.

   Single sample t-test

   Use this option when:

   - You have a single sample of scores.
   - All scores are independent from each other.

- The sampling distribution of $\bar{x}$ is normal.

In general, the single sample t-test is used to compare an average value to a known number.

Paired t-test

Use this option when:

- You have a matched pairs of scores. For example, you might have two different measures per person, or matched pairs of individuals (such as a husband and wife).
- Each pair of scores is independent of every other pair.
- The sampling distribution of $\bar{d}$ is normal.

A paired t-test is useful when comparing related cases. By averaging the differences between the scores of the paired cases, you can determine whether the total difference is statistically significant.

Unpaired t-test

Use this option when:

- You have two independent samples of scores. That is, there is no basis for pairing scores in sample 1 with those in sample 2.
- All scores within a sample are independent of all other scores within that sample.
- The sampling distribution of $x_1 - x_2$ is normal.
- Optionally, satisfy the requirement that the variance among the groups be roughly equal.

2. Set parameters dictated by the type of data you have and the type of t-test you chose. See the Options section for more information.
3. Run the experiment.
4. The output of the module is a dataset containing the t-test scores, and a transformation that you can optionally save to re-apply to this or another dataset using Apply Transformation (https://msdn.microsoft.com/en-us/library/azure/dn913055.aspx).

   No matter which type of t-test you choose, the module outputs two scores:

   - A probability score that indicates the confidence of the null hypothesis
   - A value that indicates whether the Null hypothesis should be rejected

     The goal is only to determine if we can reject the null hypothesis. Therefore, a score of 0 doesn't mean we should accept it. It means that we don't have enough data, and further investigation is needed.

Options

The options for configuring the **Hypothesis Testing T-Test** module depend on which type of t-test you apply.

Single Sample Set

Null hypothesized μ
Type the value to use as the null-hypothesized mean for the sample.

In this case, specify the expected mean value against which the sample mean will be tested.

Target column
Use the Column Selector to choose a single numeric column for testing.

Hypothesis type
Choose the type of hypothesis to use when testing. The default is a two-tailed test.

Two Tail (two-tailed test):

This is the most common type of test, in which the expected distribution is symmetric around zero.

One Tail GT (one-tailed greater than test):

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

One Tail LT (one-tailed less than test):

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

α
Specify a confidence factor.

If the value of *P* (the first output of the module) is lower than the confidence factor, the null hypothesis is rejected.

Paired Sample

Null hypothesized μ
Type a value that represents the sample difference between the pair of samples.

Target column

Use the Column Selector to choose a pair of numeric columns for testing.

Hypothesis type

Choose the type of hypothesis to use when testing. The default is a two-tailed test.

Two Tail. Use a two-tailed test.

This is the most common type of test, in which the expected distribution is symmetric around zero.

One Tail GT. Use a one-tailed greater than test.

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

One Tail LT. Use a one-tailed less than test.

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

α

Specify a confidence factor.

If the value of *P* (the first output of the module) is lower than the confidence factor, the null hypothesis is rejected.

Unpaired T-Test

Assume equal variances

Deselect this option if you want the test to

Null hypothesized μ1

Type the value to use as the null-hypothesized mean for the first sample.

In other words, provide the expected mean value against which the sample mean will be tested.

Null hypothesized μ2

Type the value to use as the null-hypothesized mean for the second sample.

In other words, provide the expected mean value against which the sample mean will be tested.

Target columns

Use the Column Selector to choose numeric columns for testing.

Hypothesis type

Choose the type of hypothesis to use when testing. The default is a two-tailed test.

Two Tail. Use a two-tailed test.

This is the most common type of test, in which the expected distribution is symmetric around zero.

One Tail GT. Use a one-tailed greater than test.

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

One Tail LT. Use a one-tailed less than test.

A one-tailed test provides more power to detect an effect in one direction, by not testing the effect in the other direction.

α

Specify a confidence factor.

If the value of *P* (the first output of the module) is lower than the confidence factor, the null hypothesis is rejected.

Technical Notes

The module automatically names the output columns according to the following conventions, depending on which type of t-test was selected, and whether the result was to reject or accept the null hypothesis.

Given input columns with names {0} and {1}, the module creates the following names:

Columns	**SingleSampleSet**	**PairedSamples**	**UnpairedSamples**
Output column P	P_ss({0})	P_ps({0}, {1})	P_us({0}, {1})
Output column **RejectH0**	RejectH0_ss({0})"	RejectH0_ps({0}, {1})	RejectH0_us({0}, {1})

How Scores are Computed

This module computes and uses the sample standard deviation; therefore, the equation used contains *(n-1)* in the denominator.

Single Sample Set

Given a single sample of scores, all independent of each other, and a normal distribution, the score is calculated as follows:

1. Take the following input:
    - A single column of values from the dataset
    - The null hypothesis (H0) parameter μ_0
    - The confidence score specified by **α**
2. Extract the number of samples (*n*).
3. Calculate the mean of the sample data.
4. Calculate the standard deviation (*s*) of the sample data.
5. Calculate *t* and degrees of freedom (*df*):

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \; df = n - 1$$

6. Extract probability *P* from distribution table *T* by using *t* and *df*.

Paired T-Test

Given a matched set of scores, with each pair independent of the other, and a normal distribution in each set, the score is calculated as follows:

1. Take the following input:
    - Two columns of values from the dataset
    - The null hypothesis (H0) parameter d_0
    - The confidence score specified by **α**
2. Extract some number of sample pairs (*n*).
3. Calculate the mean of differences for the sample data:

$$(\bar{d} = \overline{x_1 - x_2})$$

4. Calculate the standard deviation of differences (s_d).

5. Calculate *t* and the degrees of freedom (*df*):

$$t = \frac{\bar{d} - d_0}{s_d / \sqrt{n}}, \; df = n - 1$$

6. Extract probability (*P*) from the distribution table (*T*) by using *t* and *df*.

Unpaired T-Test

Given two independent samples of scores, with a normal distribution of values in each sample, the score is calculated as follows:

1. Take the following input:
    - A dataset that contains two columns of **doubles**
    - The null hypothesis (H0) parameter (d_0)
    - The confidence score specified by **α**
2. Extract a number of samples in each group, n_1 and n_2.
3. Calculate the means for each of the sample sets.
4. Calculate the standard deviation for each group as s_1 and s_2.
5. Calculate *t* and degrees of freedom (*df*):

    Optionally, satisfy the requirement that the variance among the groups be roughly equal, as follows:

    a. Calculate the pooled standard deviation first:

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

    b. If there is no assumption about variance equality, calculate as follows:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad df = n_1 + n_2 - 2$$

6. Extract *P* from the distribution table (*T*) by using *t* and *df*.

Computing Null Hypothesis

The probability of the null hypothesis, designated as *P*, is calculated as follows:

- If $P < \alpha$, set the Reject flag to True.

  If $P \geq \alpha$, set the Reject flag to False.

Expected Input

Name	Type	Description
Dataset	Data Table (https://msdn.microsoft.com/en-us/library/azure/dn905851.aspx)	Input dataset

Module Parameters

Name	Range	Type	Default	Description
Hypothesis type	Any	Hypothesis	Two-tail	Student's t-test null hypothesis type
Null hypothesized μ	Any	Float	0.0	For the single sample t-test, the null-hypothesized mean for the sample

				For the paired t-test, the sample difference
Target column(s)	Any	ColumnSelection	None	Target column(s) selection pattern
Assume equal variances	Any	Boolean	True	Assume variances of two samples are equal Applies only to unpaired samples
Null hypothesized μ1	Any	Float	0.0	Null-hypothesized mean for first sample
α	[0.0;1.0]	Float	0.95	Confidence factor (if P is lower than the confidence factor, null hypothesis is rejected)

Outputs

Name	Type	Description
P	Data Table (https://msdn.microsoft.com/en-us/library/azure/dn905851.aspx)	A probability score that indicates the confidence of the null hypothesis
Reject H0	Data Table (https://msdn.microsoft.com/en-us/library/azure/dn905851.aspx)	Value that indicates whether the Null hypothesis should be rejected

Exceptions

For a complete list of error messages, see Machine Learning Module Error Codes (https://msdn.microsoft.com/en-us/library/azure/dn905910.aspx).

Exception	Description
Error 0003 (https://msdn.microsoft.com/en-us/library/azure/dn906003.aspx)	Exception occurs if one or more of inputs are null or empty.

Error 0008 (https://msdn.microsoft.com/en-us/library/azure/dn905856.aspx)	Exception occurs if parameter is not in range.
Error 0017 (https://msdn.microsoft.com/en-us/library/azure/dn906039.aspx)	Exception occurs if one or more specified columns have a type that is unsupported by the current module.
Error 0020 (https://msdn.microsoft.com/en-us/library/azure/dn906040.aspx)	Exception occurs if the number of columns in some of the datasets passed to the module is too small.
Error 0021 (https://msdn.microsoft.com/en-us/library/azure/dn905802.aspx)	Exception occurs if the number of rows in some of the datasets passed to the module is too small.
Error 0031 (https://msdn.microsoft.com/en-us/library/azure/dn905832.aspx)	Exception occurs if the number of columns in column set is less than needed.
Error 0032 (https://msdn.microsoft.com/en-us/library/azure/dn905829.aspx)	Exception occurs if the argument is not a number.
Error 0033 (https://msdn.microsoft.com/en-us/library/azure/dn905872.aspx)	Exception occurs if the argument is infinity.

See Also

Statistical Functions (https://msdn.microsoft.com/en-us/library/azure/dn905867.aspx)
A-Z List of Machine Learning Studio Modules (https://msdn.microsoft.com/en-us/library/azure/dn906033.aspx)