ttest

One-sample and paired-sample t-test

Syntax

h = ttest(x)

h = ttest(x,m)

h = ttest(x,y)

h = ttest(...,alpha)

h = ttest(...,alpha,tail)

h = ttest(...,alpha,tail,dim)

[h,p] = ttest(...)

[h,p,ci] = ttest(...)

[h,p,ci,stats] = ttest(...)

Description

h = ttest(x) performs a t-test of the null hypothesis that data in the vector x are a random sample from a normal distribution with mean 0 and unknown variance, against the alternative that the mean is not 0. The result of the test is returned in h. h = 1 indicates a rejection of the null hypothesis at the 5% significance level. h = 0 indicates a failure to reject the null hypothesis at the 5% significance level.

x can also be a matrix or an N-dimensional array. For matrices, ttest performs separate t-tests along each column of x and returns a vector of results. For N-dimensional arrays, ttest works along the first non-singleton dimension of x.

The test treats NaN values as missing data, and ignores them.

h = ttest(x,m) performs a t-test of the null hypothesis that data in the vector x are a random sample from a normal distribution with mean m and unknown variance, against the alternative that the mean is not m.

h = ttest(x,y) performs a paired t-test of the null hypothesis that data in the difference x-y are a random sample from a normal distribution with mean 0 and unknown variance, against the alternative that the mean is not 0. x and y must be vectors of the same length, or arrays of the same size.

h = ttest(...,alpha) performs the test at the (100\*alpha)% significance level. The default, when unspecified, is alpha = 0.05.

h = ttest(...,alpha,tail) performs the test against the alternative specified by the string tail. There are three options for tail:

'both' — Mean is not 0 (or m) (two-tailed test). This is the default, when tail is unspecified.

'right' — Mean is greater than 0 (or m) (right-tail test)

'left' — Mean is less than 0 (or m) (left-tail test)

tail must be a single string, even when x is a matrix or an N-dimensional array.

h = ttest(...,alpha,tail,dim) works along dimension dim of x, or of x-y for a paired test. Use [] to pass in default values for m, alpha, or tail.

[h,p] = ttest(...) returns the p value of the test. The p value is the probability, under the null hypothesis, of observing a value as extreme or more extreme of the test statistic

where is the sample mean, μ = 0 (or m) is the hypothesized population mean, s is the sample standard deviation, and n is the sample size. Under the null hypothesis, the test statistic will have Student's t distribution with n – 1 degrees of freedom.

[h,p,ci] = ttest(...) returns a 100\*(1 – alpha)% confidence interval on the population mean, or on the difference of population means for a paired test.

[h,p,ci,stats] = ttest(...) returns the structure stats with the following fields:

tstat — Value of the test statistic

df — Degrees of freedom of the test

sd — Sample standard deviation

Examples

Simulate a random sample of size 100 from a normal distribution with mean 0.1:

x = normrnd(0.1,1,1,100);

Test the null hypothesis that the sample comes from a normal distribution with mean 0:

[h,p,ci] = ttest(x,0)

h =

 0

p =

 0.8323

ci =

 -0.1650 0.2045

The test fails to reject the null hypothesis at the default α = 0.05 significance level. Under the null hypothesis, the probability of observing a value as extreme or more extreme of the test statistic, as indicated by the p value, is much greater than α. The 95% confidence interval on the mean contains 0.

Simulate a larger random sample of size 1000 from the same distribution:

y = normrnd(0.1,1,1,1000);

Test again if the sample comes from a normal distribution with mean 0:

[h,p,ci] = ttest(y,0)

h =

 1

p =

 0.0160

ci =

 0.0142 0.1379

This time the test rejects the null hypothesis at the default α = 0.05 significance level. The p value has fallen below α = 0.05 and the 95% confidence interval on the mean does not contain 0.

Because the p value of the sample y is greater than 0.01, the test will fail to reject the null hypothesis when the significance level is lowered to α = 0.01:

[h,p,ci] = ttest(y,0,0.01)

h =

 0

p =

 0.0160

ci =

 -0.0053 0.1574

Notice that at the lowered significance level the 99% confidence interval on the mean widens to contain 0.

This example will produce slightly different results each time it is run, because of the random sampling.

ttest2

Two-sample t-test

Syntax

h = ttest2(x,y)

h = ttest2(x,y,alpha)

h = ttest2(x,y,alpha,tail)

h = ttest2(x,y,alpha,tail,vartype)

h = ttest2(x,y,alpha,tail,vartype,dim)

[h,p] = ttest2(...)

[h,p,ci] = ttest2(...)

[h,p,ci,stats] = ttest2(...)

Description

h = ttest2(x,y) performs a t-test of the null hypothesis that data in the vectors x and y are independent random samples from normal distributions with equal means and equal but unknown variances, against the alternative that the means are not equal. The result of the test is returned in h. h = 1 indicates a rejection of the null hypothesis at the 5% significance level. h = 0 indicates a failure to reject the null hypothesis at the 5% significance level. x and y need not be vectors of the same length.

x and y can also be matrices or N-dimensional arrays. Matrices x and y must have the same number of columns, in which case ttest2 performs separate t-tests along each column and returns a vector of results. N-dimensional arrays x and y must have the same size along all but the first non-singleton dimension, in which case ttest2 works along the first non-singleton dimension.

The test treats NaN values as missing data, and ignores them.

h = ttest2(x,y,alpha) performs the test at the (100\*alpha)% significance level. The default, when unspecified, is alpha = 0.05.

h = ttest2(x,y,alpha,tail) performs the test against the alternative specified by the string tail. There are three options for tail:

'both' — Means are not equal (two-tailed test). This is the default, when tail is unspecified.

'right' — Mean of x is greater than mean of y (right-tail test)

'left' — Mean of x is less than mean of y (left-tail test)

tail must be a single string, even when x is a matrix or an N-dimensional array.

h = ttest2(x,y,alpha,tail,vartype) performs the test under the assumption of equal or unequal population variances, as specified by the string vartype. There are two options for vartype:

'equal' — Assumes equal variances. This is the default, when vartype is unspecified.

'unequal' — Does not assume equal variances. This is the Behrens-Fisher problem.

vartype must be a single string, even when x is a matrix or an N-dimensional array.

If vartype is 'equal', the test computes a pooled sample standard deviation using

where sx and sy are the sample standard deviations of x and y, respectively, and n and m are the sample sizes of x and y, respectively.

h = ttest2(x,y,alpha,tail,vartype,dim) works along dimension dim of x and y. Use [] to pass in default values for alpha, tail, or vartype.

[h,p] = ttest2(...) returns the p value of the test. The p value is the probability, under the null hypothesis, of observing a value as extreme or more extreme of the test statistic

where and are the sample means, sx and sy are the sample standard deviations (replaced by the pooled standard deviation s in the default case where vartype is 'equal'), and n and m are the sample sizes.

In the default case where vartype is 'equal', the test statistic, under the null hypothesis, has Student's t distribution with n + m – 2 degrees of freedom.

In the case where vartype is 'unequal', the test statistic, under the null hypothesis, has an approximate Student's t distribution with a number of degrees of freedom given by Satterthwaite's approximation.

[h,p,ci] = ttest2(...) returns a 100\*(1 – alpha)% confidence interval on the difference of population means.

[h,p,ci,stats] = ttest2(...) returns structure stats with the following fields:

tstat — Value of the test statistic

df — Degrees of freedom of the test

sd — Pooled sample standard deviation (in the default case where vartype is 'equal') or a vector with the sample standard deviations (in the case where vartype is 'unequal').

Examples

Simulate random samples of size 1000 from normal distributions with means 0 and 0.1, respectively, and standard deviations 1 and 2, respectively:

x = normrnd(0,1,1,1000);

y = normrnd(0.1,2,1,1000);

Test the null hypothesis that the samples come from populations with equal means, against the alternative that the means are unequal. Perform the test assuming unequal variances:

[h,p,ci] = ttest2(x,y,[],[],'unequal')

h =

 1

p =

 0.0102

ci =

 -0.3227 -0.0435

The test rejects the null hypothesis at the default α = 0.05 significance level. Under the null hypothesis, the probability of observing a value as extreme or more extreme of the test statistic, as indicated by the p value, is less than α. The 95% confidence interval on the mean of the difference does not contain 0.

This example will produce slightly different results each time it is run, because of the random sampling.

ztest

z-test

Syntax

h = ztest(x,m,sigma)

h = ztest(...,alpha)

h = ztest(...,alpha,tail)

h = ztest(...,alpha,tail,dim)

[h,p] = ztest(...)

[h,p,ci] = ztest(...)

[h,p,ci,zval] = ztest(...)

Description

h = ztest(x,m,sigma) performs a z-test of the null hypothesis that data in the vector x are a random sample from a normal distribution with mean m and standard deviation sigma, against the alternative that the mean is not m. The result of the test is returned in h. h = 1 indicates a rejection of the null hypothesis at the 5% significance level. h = 0 indicates a failure to reject the null hypothesis at the 5% significance level.

x can also be a matrix or an N-dimensional array. For matrices, ztest performs separate z-tests along each column of x and returns a vector of results. For N-dimensional arrays, ztest works along the first non-singleton dimension of x.

The test treats NaN values as missing data, and ignores them.

h = ztest(...,alpha) performs the test at the (100\*alpha)% significance level. The default, when unspecified, is alpha = 0.05.

h = ztest(...,alpha,tail) performs the test against the alternative specified by the string tail. There are three options for tail:

'both' — Mean is not m (two-tailed test). This is the default, when tail is unspecified.

'right' — Mean is greater than m (right-tail test)

'left' — Mean is less than m (left-tail test)

tail must be a single string, even when x is a matrix or an N-dimensional array.

h = ztest(...,alpha,tail,dim) works along dimension dim of x. Use [] to pass in default values for alpha or tail.

[h,p] = ztest(...) returns the p value of the test. The p value is the probability, under the null hypothesis, of observing a value as extreme or more extreme of the test statistic

where is the sample mean, μ = m is the hypothesized population mean, σ is the population standard deviation, and n is the sample size. Under the null hypothesis, the test statistic will have a standard normal distribution, N(0,1).

[h,p,ci] = ztest(...) returns a 100\*(1 – alpha)% confidence interval on the population mean.

[h,p,ci,zval] = ztest(...) returns the value of the test statistic.

Examples

Simulate a random sample of size 100 from a normal distribution with mean 0.1 and standard deviation 1:

x = normrnd(0.1,1,1,100);

Test the null hypothesis that the sample comes from a standard normal distribution:

[h,p,ci] = ztest(x,0,1)

h =

 0

p =

 0.1391

ci =

 -0.0481 0.3439

The test fails to reject the null hypothesis at the default α = 0.05 significance level. Under the null hypothesis, the probability of observing a value as extreme or more extreme of the test statistic, as indicated by the p value, is greater than α. The 95% confidence interval on the mean contains 0.

Simulate a larger random sample of size 1000 from the same distribution:

y = normrnd(0.1,1,1,1000);

Test again if the sample comes from a normal distribution with mean 0:

[h,p,ci] = ztest(y,0,1)

h =

 1

p =

 5.5160e-005

ci =

 0.0655 0.1895

This time the test rejects the null hypothesis at the default α = 0.05 significance level. The p value has fallen below α = 0.05 and the 95% confidence interval on the mean does not contain 0.

Because the p value of the sample y is less than 0.01, the test will still reject the null hypothesis when the significance level is lowered to α = 0.01:

[h,p,ci] = ztest(y,0,1,0.01)

h =

 1

p =

 5.5160e-005

ci =

 0.0461 0.2090