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

$$s = \sqrt{\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}$$

where s_x and s_y 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

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_x^2}{n} + \dfrac{s_y^2}{m}}}$$

where $\bar{x}$ and $\bar{y}$ are the sample means, s_x and s_y are the sample standard deviations (replaced by the pooled standard deviations 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 $\alpha = 0.05$ 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.

See Also

ttest, ztest

Was this topic helpful?   Yes   No