

```

> A_c(13.2,8.2,10.9,14.3,10.7,6.6,9.5,10.8,8.8,13.3)
> B_c(14,8.8,11.2,14.2,11.8,6.4,9.8,11.3,9.3,13.6)
> shoes_data.frame(A,B)
> shoes
  A   B
1 13.2 14.0
2  8.2  8.8
3 10.9 11.2
4 14.3 14.2
5 10.7 11.8
6  6.6  6.4
7  9.5  9.8
8 10.8 11.3
9  8.8  9.3
10 13.3 13.6

> t.test(shoes$A, mu=10)

```

One-sample t-Test

```

data: shoes$A
t = 0.8127, df = 9, p-value = 0.4373
alternative hypothesis: true mean is not equal to 10
95 percent confidence interval:
 8.876427 12.383573
sample estimates:
mean of x
 10.63

> attach(shoes)
> t.test(A, mu=10)

```

One-sample t-Test

```

data: A
t = 0.8127, df = 9, p-value = 0.4373
alternative hypothesis: true mean is not equal to 10
95 percent confidence interval:
 8.876427 12.383573
sample estimates:
mean of x
 10.63
> t.test(A)$conf.int
[1] 8.876427 12.383573
attr(,"conf.level"):
[1] 0.95
> wilcox.test(A,mu=10)

```

Exact Wilcoxon signed-rank test

data: A
signed-rank statistic $V = 34$, $n = 10$, $p\text{-value} = 0.5566$
alternative hypothesis: true mu is not equal to 10

```
>help(wilcox.test)
```

DESCRIPTION

Computes Wilcoxon rank sum test for two sample data (equivalent to the Mann-Whitney test) or the Wilcoxon signed rank test for paired or one sample data.

USAGE

```
wilcox.test(x, y, alternative="two.sided", mu=0, paired=F, exact=T,  
            correct=T)
```

REQUIRED ARGUMENTS

x numeric vector. Missing values (NAs) are allowed but ignored at calculation.

OPTIONAL ARGUMENTS

y numeric vector. If supplied a two sample test is performed. If `paired=TRUE` then `x` and `y` must have the same length. Missing values (NAs) are allowed but ignored at calculation.

alternative a character string specifying the alternative hypothesis, "greater" (the location of `x` is greater than that of `y`) or "less" (the location of `x` is less than that of `y`) or "two.sided" (the locations of `x` and `y` are different). Only enough of the string to determine a unique match is required.

mu the location shift for the distribution of `x`.

paired if `TRUE`, the Wilcoxon signed rank test is computed. The default is the Wilcoxon rank sum test.

exact if `TRUE` the exact distribution for the test statistic is used to compute the p -value if possible.

correct if `TRUE` a continuity correction is applied to the normal approximation for the p -value.

VALUE

a list with class attribute "htest", representing the result of the Wilcoxon rank sum or Wilcoxon signed rank test:

statistic the test statistic with a name describing the statistic.

parameters the parameters for the test statistic with names attribute "n" (number of meaningful observations in `x`) for the exact Wilcoxon signed rank statistic or "n" and "m" (number of meaningful observations in `x` and `y`, respectively) for the exact Wilcoxon rank sum test. This component is missing if a normal approximation is used for the test.

p.value	the p-value for the test.
null.value	the value of the location values or difference in locations specified by the null hypothesis. This equals the input argument mu with names attribute "mu".
alternative	the input argument alternative.
method	the name of the test applied.
data.name	a character string describing the actual data names used for the test.

NULL HYPOTHESIS

For the Wilcoxon rank sum test, the null hypothesis is that the locations of the distributions of x and y differ by μ .

For the one-sample Wilcoxon signed rank test, the null hypothesis is that the median of the distribution of x (if only x is supplied) is μ . For the two-sample paired case, where both x and y are provided (and paired is TRUE), the null hypothesis is that the median of the distribution of $x - y$ is μ .

ASSUMPTIONS

For the Wilcoxon rank sum test, the x and y vectors are independent samples from their respective distributions, and there is mutual independence between the two sample.

For the Wilcoxon signed rank test the values of x or $x - y$ are independent observations from the same symmetric distribution.

DETAILS

The Wilcoxon rank sum test statistic, W , is defined as $\text{sum}(\text{rank}(c(x - \mu, y))[\text{seq}(\text{along}=x)])$. To compute probabilities for large datasets, a normal approximation is used for the distribution of the test statistic when the length of x or y is greater than 49 (or when `exact=F`). See the hard copy help file for the Normal approximation used for W when the length of x or y is greater than 49. If there are ties in the vector $c(x - \mu, y)$ an exact p-value for the test cannot be computed. In this case, the normal approximation given by Lehmann (1975, p. 20) is used.

The Wilcoxon signed-rank test statistic, V , is the rank sum of those $\text{abs}(x - \mu - y)$ with $x - \mu - y > 0$. For $n = \text{length}(x)$ greater than 25, a normal approximation is used to compute the probability (or if `exact=F`). See the hard copy help file for the Normal approximation used for the test statistic. If there are ties in the absolute values of the differences, $x - \mu - y$, an exact probability for the test cannot be computed. The normal approximation given by Lehmann (1975, p. 130) is used in this case.

REFERENCES

Conover, W. J. (1980). Practical Nonparametric Statistics, 2nd ed. Wiley, New York.

Lehmann, E. L. (1975). Nonparametrics: Statistical Methods Based on Ranks. Holden and Day, San Francisco.

SEE ALSO

Wilcoxon, t.test.

EXAMPLES

```
# A Wilcoxon rank sum test
```

```
wilcox.test(x, y)
```

```
# A Wilcoxon signed rank sum test
```

```
wilcox.test(x, y, paired=T)
```

```
> var.test(A,B)
```

F test for variance equality

```
data: A and B
```

```
F = 0.9474, num df = 9, denom df = 9, p-value = 0.9372
```

```
alternative hypothesis: true ratio of variances is not equal to 1
```

```
95 percent confidence interval:
```

```
0.2353191 3.8142000
```

```
sample estimates:
```

```
variance of x variance of y
```

```
6.009 6.342667
```

```
> help(var.test)
```

```
> t.test(A,B)
```

Standard Two-Sample t-Test

```
data: A and B
```

```
t = -0.3689, df = 18, p-value = 0.7165
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-2.744924 1.924924
```

```
sample estimates:
```

```
mean of x mean of y
```

```
10.63 11.04
```

DESCRIPTION

Performs a one-sample, two-sample, or paired t-test, or a Welch modified two-sample t-test.

USAGE

```
t.test(x, y=NULL, alternative="two.sided", mu=0, paired=F,  
var.equal=T, conf.level=.95)
```

NULL HYPOTHESIS

For the one-sample t-test, the null hypothesis is that the mean of the population from which x is drawn is μ . For the paired t-test, the null hypothesis is that the population mean of the difference $x - y$ is equal to μ . For the standard and Welch modified two-sample t-tests, the null hypothesis is that the population mean for x is less than that for y is μ .

The alternative hypothesis in each case indicates the direction of divergence of the population mean for x
(or difference of means for x and y) from μ (i.e., "greater", "less", "two.sided").

TEST ASSUMPTIONS

The assumption of equal population variances is central to the standard two-sample t-test. This test can be misleading when population variances are not equal, as the null distribution of the test statistic is no longer a t-distribution. If the assumption of equal variances is doubtful with respect to a particular dataset, the Welch modification of the t-test should be used.

The t-test and the associated confidence interval are quite robust with respect to level toward heavy-tailed non-Gaussian distributions (e.g., data with outliers). However, the t-test is quite non-robust with respect to power, and the confidence interval is quite non-robust with respect to average length, toward these same types of distributions.

DETAILS

(a) One-Sample t-Test.

The arguments `y`, `paired` and `var.equal` determine the type of test. If `y` is `NULL`, a one-sample t-test is carried out with `x`. Here statistic is given by:

$$t \leftarrow (\text{mean}(x) - \mu) / (\sqrt{\text{var}(x)} / \sqrt{\text{length}(x)})$$

If `x` was drawn from a normal population, `t` has a t-distribution with $\text{length}(x) - 1$ degrees of freedom under the null hypothesis.

(b) Paired t-Test.

If `y` is not `NULL` and `paired=TRUE`, a paired t-test is performed; here statistic is defined through

$$t \leftarrow (\text{mean}(d) - \mu) / (\sqrt{\text{var}(d)} / \sqrt{\text{length}(d)})$$

where `d` is the vector of differences $x - y$. Under the null hypothesis, `t` follows a t-distribution with $\text{length}(d) - 1$ degrees of freedom, assuming normality of the differences `d`.

(c) Standard Two-Sample t-Test.

If y is not NULL and paired=FALSE, either a standard or Welch modified two-sample t-test is performed, depending on whether var.equal is TRUE or FALSE. For the standard t-test, statistic is

```
t <- (mean(x) - mean(y) - mu) / s1,
```

with

```
s1 <- sp * sqrt(1/nx + 1/ny),
```

```
sp <- sqrt( ((nx-1)*var(x) + (ny-1)*var(y)) / (nx + ny - 2) ),
```

```
nx <- length(x), ny <- length(y).
```

Assuming that x and y come from normal populations with equal variances, t has a t-distribution with $nx + ny - 2$ degrees of freedom under the null hypothesis.

(d) Welch Modified Two-Sample t-Test.

If y is not NULL, paired=FALSE and var.equal=FALSE, the Welch modified two-sample t-test is performed. In this case statistic is

```
t <- (mean(x) - mean(y) - mu) / s2
```

with

```
s2 <- sqrt( var(x)/nx + var(y)/ny ),
```

```
nx <- length(x), ny <- length(y).
```

If x and y come from normal populations, the distribution of t under the null hypothesis can be approximated by a t-distribution with (non-integral) degrees of freedom

$$1 / ((c^2)/(nx-1) + ((1-c)^2)/(ny-1))$$

where

```
> t.test(A,B)
```

Standard Two-Sample t-Test

data: A and B

t = -0.3689, df = 18, p-value = 0.7165

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.744924 1.924924

sample estimates:

mean of x mean of y

10.63 11.04

```
> help(t.test)
> t.test(A,B,paired=T)
```

Paired t-Test

```
data: A and B
t = -3.3489, df = 9, p-value = 0.0085
alternative hypothesis: true mean of differences is not equal to 0
95 percent confidence interval:
-0.6869539 -0.1330461
sample estimates:
mean of x - y
-0.41
> t.test(A,B,var.equal=F)
```

Welch Modified Two-Sample t-Test

```
data: A and B
t = -0.3689, df = 17.987, p-value = 0.7165
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2.745046 1.925046
sample estimates:
mean of x mean of y
10.63 11.04
> wilcox.test(A,B)
```

Wilcoxon rank-sum test

```
data: A and B
rank-sum normal statistic with correction Z = -0.5293, p-value = 0.5966
alternative hypothesis: true mu is not equal to 0
```

Warning messages:

```
cannot compute exact p-value with ties in: wil.rank.sum(x, y, alternat
ive, exact, correct)
> wilcox.test(A,B,paired=T)
```

Wilcoxon signed-rank test

```
data: A and B
signed-rank normal statistic with correction Z = -2.4495, p-value = 0.01
43
alternative hypothesis: true mu is not equal to 0
```

Warning messages:

```
cannot compute exact p-value with ties in: wil.sign.rank(dff, alternat
ive, exact, correct)
```

```

> eda.shape_function(x) {
+   par(mfrow=c(2,2))
+   hist(x)
+   boxplot(x)
+   iqd_summary(x)[5]-summary(x)[2]
+   plot(density(x,width=2*iqd),xlab="x",ylab="",type="l")
+   qqnorm(x)
+   qqline(x)
+   par(mfrow=c(1,1))
+   invisible()
+ }
> eda.shape(B-A)

```

