Package 'clusrank'

October 6, 2016

Date 2016-9-25

Version 0.4-5

Title Wilcoxon Rank Sum Test for Clustered Data

Description
Non-parametric tests (Wilcoxon rank sum test and Wilcoxon signed rank test) for clustered data.

Imports stats, MASS, Rcpp (>= 0.12.2)

License GPL (>= 3)

Depends R (>= 3.2.0)

LinkingTo Rcpp

LazyData true

RoxygenNote 5.0.1

NeedsCompilation yes

Author Yujing Jiang [aut, cre],
Mei-Ling Ting Lee [ctb],
Jun Yan [ctb]

Maintainer Yujing Jiang <yujing.jiang@uconn.edu>

Repository CRAN

Date/Publication 2016-10-06 13:06:08

R topics documented:

clusWilcox.test *Wilcoxon Rank Sum and Signed Rank Test for Clustered Data*

Description

Performs one-sample and two-sample Wilcoxon test for clutered data on vectors of data.

Usage

```
clusWilcox.test(x, ...)

## S3 method for class 'formula'
clusWilcox.test(formula, data = parent.frame(),
  subset = NULL, na.action = na.omit, alternative = c("two.sided", "less",
  "greater"), mu = 0, paired = FALSE, exact = FALSE, method = c("rgl",
  "ds"), ...)

## Default S3 method:
clusWilcox.test(x, y = NULL, cluster = NULL,
  group = NULL, stratum = NULL, data = NULL,
  alternative = c("two.sided", "less", "greater"), mu = 0, paired = FALSE,
  exact = FALSE, method = c("rgl", "ds"), ...)
```

Arguments

x	A numeric vector of data values or a formula. Non-finite (e.g., infinite or missing) values will be omitted.
...	Further arguments to be passed to or from methods.
formula	A formula of the form `lhs ~ rhs` where the `lhs` is the measurements and the `rhs` is of the form group + `cluster`(x1) + `stratum`(x2) for clustered rank sum test, where `x1` and `x2` are cluster id and stratum id in the data frame `data`. For clustered signed rank test, the `rhs` only contains `clusterx1`.
data	An optional dataframe containing the variables.
subset	An optional vector specifying a subset of observations to be used.
na.action	A function which indicates what should happen when the data contain NAs. Defaults to `getOption("na.action")`.
alternative	A character string specifying the alternative hypothesis, must be one of `"two sided"` (default), `"greater"` or `"less"`. You can specify just the initial letter.
mu	A number specifying an optional parameter used to form the null hypothesis. Default is 0. See 'Details'.
paired	A logical indicating whether you want a paired test.
exact	A logical indicating whether an exact p-value should be computed. Only available for `rgl` signed rank test and `rgl` rank sum test when treatment is assigned at cluster leve.

method	A character string specifying the method of clustered wilcoxon rank test to be use, should be one of "rgl" or "ds".
y	An optional numeric vector of data values, non-finite values will be omitted.
cluster	An optional numeric vector of cluster id.
group	An optional numeric vector of treatment id.
stratum	An opptional numeric vector of stratum id. Only available for rgl rank sum test when treatment is assigned at cluster level.

Details

The formula interface is to both clustered signed rank test and clustered rank sum test.

The default of cluster id is that there is one member in each cluster. Both balanced data (identical cluster size) and unbalanced data (different cluster sizes) are supported in all tests provided in this package. For clustered rank sum test, the data can either have treatment assigned at cluster level or individual level.

If both x and y are given or only x is given and paired is TRUE, a clustered Wilcoxon signed rank test of the null that the distribution of x - y (paired sample) or of x (one sample) is symmetric about mu is performed.

Otherwise, if only x is given and paired is FALSE, a Wilcoxon rank sum test is performed. In this case, measurements from different treatment groups should be combined in x and the group variable is required. When there are two treatment groups, the null is that the distributions of values from the two groups differ by a location shift of mu and the alternative is that they differ by some other location shift. When there are m (>= 2) treatment groups, ds method can test if the location of the m groups are identical or not.

For RGL rank sum test when treatment is assigned at cluster level, an extra stratification variable is allowed through stratum.

The exact test is only available for RGL signed rank test and RGL rank sum test when treatment is assigned at cluster level.

Value

A list with class "htest" containing the following components, for different test the components may vary:

Rstat	the value of the rank statistic with a name describing it.
ERstat	the expectation of the rank statistic.
VRstat	the variance of the rank statistic.
statistic	the value of the test statistic with a name describing it.
p.value	the p-value for the test.
alternative	a character string describing the alternative hypothesis.
null.value	the location parameter 'mu'.
method	the type of test applied.
data.name	a character string giving the names of the data.
balance	a logical indicating whether the data set is balanced.

n.group	number of treatment, will be returned if there are more than 2 treatment groups and ds method is used.
df	degrees of freedom of chi-square distribution, will be returned when there are more than 2 treatment groups and ds method is used.
n	number of observations
cn	number of clusters

Methods (by class)

- formula: S3 method for class 'formula'
- default: Default S3 method.

Warning

This function can use large amounts of memory and stack if 'exact = TRUE' and one sample is large (and even crash R if the stack limit is exceeded). Not recommended for data set with number of clusters more than 50.

Author(s)

Yujing Jiang

References

Bernard Rosner, Robert J. Glynn, Mei-Ling T. Lee (2006) *The Wilcoxon Signed Rank Test for Paired Comparisons of Clustered Data*. Biometrics, **62**, 185-192.

Bernard Rosner, Robert J. Glynn, Mei-Ling T. Lee (2003) *Incorporation of Clustering Effects for the Wilcoxon Rank Sum Test: A Large-Sample Approach*. Biometrics, **59**, 1089-1098.

Bernard Rosner, Robert J. Glynn, Mei-Ling T. Lee (2006) *Extension of the Rank Sum Test for Clustered Data: Two-Group Comparisons with Group*. Biometrics, **62**, 1251-1259.

Somnath Datta, Glen A. Satten (2005) *Rank-Sum Tests for Clustered Data*. Journal of the American Statistical Association, **100**, 908-915.

Somath Datta, Glen A. Satten (2008) *A Signed-Rank test for Clustered Data*. Biometric, **64**, 501-507.

Examples

```
## Clustered signed rank test using RGL method.
data(crsd)
clusWilcox.test(z, cluster = id, data = crsd, paired = TRUE)
## or
clusWilcox.test(z ~ cluster(id), data = crsd, paired = TRUE)
## Not run: clusWilcox.test(z, cluster = id, data = crsd)
## Default is rank sum test. The group variable is required.
## End(Not run)
## Clustered rank sum test using RGL method.
data(crd)
clusWilcox.test(z ~ group + cluster(id), data = crd)
```

```
## or
clusWilcox.test(z, cluster = id, group = group, data = crd)
```

Index