Package 'clusrank'

October 12, 2022

Version 1.0-3

Title Wilcoxon Rank Tests for Clustered Data

Description Non-parametric tests (Wilcoxon rank sum test and Wilcoxon signed rank test) for clustered data documented in Jiang et. al (2020) <doi:10.18637/jss.v096.i06>.

License GPL (>= 3)

Depends R (>= 4.0)

Imports stats, MASS, Rcpp (>= 1.0.5)

LinkingTo Rcpp

LazyData true

URL https://github.com/wenjie2wang/clusrank

BugReports https://github.com/wenjie2wang/clusrank/issues

Encoding UTF-8

RoxygenNote 7.2.1

NeedsCompilation yes

Author Wenjie Wang [cre, ctb] (<https://orcid.org/0000-0003-0363-3180>),
Yujing Jiang [aut],
Mei-Ling Ting Lee [ctb],
Jun Yan [ctb]

Maintainer Wenjie Wang <wang@wwenjie.org>

Repository CRAN

Date/Publication 2022-08-12 06:10:01 UTC

R topics documented:

---

amd   *CARMS scores*

Description

A data set from a research on complement factor H R1210C rare variant and its associated pheno-type. This data set contains Clinical Age-Related Maculopathy Staging (CARMS) scores from a total of 143 patients (283 eyes), including 62 patients with the rare variant, The data is from the lab of Dr. Johanna M. Seddon,

Format

A data frame with 283 rows and 7 variables.

- ID patient identifier
- Eye OD (right eye), OS(left eye)
- Variant 1: No R1210C Variant; 2: R1210C Variant
- CARMS Patient's last CARMS grade, related to age-related macular degeneration (AMD). 1: no AMD; 2: early AMD; 3: intermediate AMD; 4: geographic atroph(advanced dry); 5: neovascular disease (advanced wet)
- Age_group 1: < 70 years; 2: 70 to 79.9 years; 3: >= 80 years
- Sex 1: male; 2: female
- AgeSex 1: agegroup = 1, sex = 1; 2: agegroup = 2, sex = 1; 3: agegroup = 3, sex = 1; 4: agegroup = 1, sex = 2; 5: agegroup = 2, sex = 2; 6: agegroup = 3, sex = 2;

Note

CARMS grades were assessed separately for the two advanced stages (4 and 5): 1. CARMS 1,2,3, and 4 was assessed; 2. CARMS 1,2,3, and 5 was assessed

Source

The data came from Seddon's lab.

References

Seddon JM, Sharma S, Adelman RA (2006) *Evaluation of the Clinical Age-related Maculopathy Staging System.* Ophthalmology, **113**, 260-266.

Ferrara D, Seddon JM (2015) *Phenotypic characterization of complement factor H R1210C rare genetic variant in age-related macular degeneration* JAMA Ophthalmol, 2015 Apr 16. doi:10.1001/jamaophthalmol.2015.0814

Examples

```
data(amd)
clusWilcox.test(CARMS ~ Variant + cluster(ID), data = amd,
                subset = CARMS %in% c(1, 2, 3, 4), method = "rgl", alternative = "two")
clusWilcox.test(CARMS ~ Variant + cluster(ID), data = amd,
                subset = CARMS %in% c(1, 2, 3, 4), method = "ds", alternative = "two")
clusWilcox.test(CARMS ~ Variant + cluster(ID) + stratum(AgeSex), data = amd,
                subset = CARMS %in% c(1, 2, 3, 4), alternative = "two")
```

---

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

Description

Performs one-sample and two-sample Wilcoxon test for clustered 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,
  B = 2000,
  method = c("rgl", "ds", "dd"),
  ...
)

## 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,
```

```
  B = 2000,
  method = c("rgl", "ds", "dd"),
  ...
)
```

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 should only contain cluster(x1).
data	An optional data frame 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 if a permutation test is to be used. If it is set as FALSE, then test based on large-sample theory is carried out. If it is set as TRUE, then a permutation test is carried out. There are two available permutation tests: exact permutation test is available for rgl signed rank test and rgl rank sum test when treatment is assigned at cluster level, random permutation test is available for all tests. For more details look at the description of the argument B.
B	An integer specifying the number of permutations if exact = TRUE. If B = 0, then the test is the exact permutation test. If B > 0, then an approximation to the exact test with B random permutation. Default is set as 2000 to ran a random permutation test.
method	A character string specifying the method of clustered Wilcoxon rank test to be use, should be one of "rgl" or "ds". For rank-sum test, the "dd" method is also available.
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 optional 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.
ngroup	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.
nobs	number of observations
nclus	number of clusters

Methods (by class)

- `clusWilcox.test(formula)`: S3 method for class 'formula'

- `clusWilcox.test(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*. Biometrics, **64**, 501-507.

Sandipan Dutta, Somnath Datta (2015) *A Rank-Sum Test for Clustered Data When the Number of Subjects in a Group within a Cluster is Informative*. Biometrics, **72**, 432-440.

Examples

```
## Clustered signed rank test using RGL method.
data(crsd)
clusWilcox.test(z, cluster = id, data = crsd, paired = TRUE, method = "rgl")
## or
clusWilcox.test(z ~ cluster(id), data = crsd, paired = TRUE, method = "rgl")
## 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, method = "rgl")
## or
clusWilcox.test(z, cluster = id, group = group, data = crd, method = "rgl")
```

---

pairwise.clusWilcox.test

Pairwise Wilcoxon Rank Sum and Signed Rank Tests for Clustered Data

---

Description

Performs pairwise comparisons between group levels with corrections for multiple testing.

Usage

```
pairwise.clusWilcox.test(
  x,
  group,
  cluster,
  p.adjust.method = p.adjust.methods,
  ...
)
```

Arguments

x	A numeric vector of data values.
group	An optional numeric vector of treatment id.
cluster	An optional numeric vector of cluster id.
p.adjust.method	
	Method for adjusting p values. ?p.adjust for details.
...	Other arguments passed to clusWilcox.test.

Value

A pairwise.htest object.

Examples

```
library(clusrank)
data(crd)

## for demonstration purpose, create random groups
set.seed(123)
g <- sample(seq_len(3), size = nrow(crd), replace = TRUE)
with(crd, pairwise.clusWilcox.test(z, group = g, cluster = id,
                                   method = "ds", p.adjust.method = "fdr"))
```

Index