Package 'npclust'

July 22, 2025

Type Package

Title Nonparametric Tests for Incomplete Clustered Data

Version 0.1.1

Date 2024-12-03

Maintainer Yue Cui <YueCui@MissouriState.edu>

Description Nonparametric tests for clustered data in pre-post intervention design documented in Cui and Harrar (2021) <doi:10.1002/bimj.201900310> and Harrar and Cui (2022) <doi:10.1016/j.jspi.2022.05.009>. Other than the main test results mentioned in the reference paper, this package also provides a function to calculate the sample size allocations for the input long format data set, and also a function for adjusted/unadjusted confidence intervals calculations. There are also functions to visualize the distribution of data across different intervention groups over time, and also the adjusted/unadjusted confidence intervals.

License GPL (>= 2)

Encoding UTF-8

LazyData true

Repository CRAN

Date/Publication 2024-12-03 07:20:02 UTC

RoxygenNote 7.2.3

Depends R (>= 2.10)

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Imports graphics, stats, MASS, ggplot2

NeedsCompilation no

Author Yue Cui [aut, cre] (ORCID: <https://orcid.org/0000-0002-7304-9409>), Solomon W. Harrar [aut] (ORCID: <https://orcid.org/0000-0001-6802-340X>)

Contents

`ARTIS` *Asthma Randomized Trial of Indoor-Wood Smoke Data*

Description

A subset of the data set from a randomized trial of interventions to improve childhood asthma in homes with wood-burning stoves. The original data was collected on 115 children with asthma living in 98 eligible households. The outcomes are domain scores for Pediatric Asthma Quality of Life Questionnaire (PAQLQ) in symptoms, activity limitation and emotional function.

Usage

```
ARTIS
```

Format

A data frame with 50 rows and 6 variables:

`homeid` unique id for each household

`intervention` indicator for intervention, where 0 = pre-intervention, 1=after-intervention

`tx` intervention type, where 1 = sham fiter, 2 = updated wood-burning stove, 3 = air-filter

`symptoms_pqol` PAQLQ for symtoms

`act_pqol` PAQLQ score for activity

`emot_pqol` PAQLQ score for emotional function

Source

Noonan, Curtis W., and Tony J. Ward. "Asthma randomized trial of indoor wood smoke (ARTIS): rationale and methods." Contemporary clinical trials 33, no. 5 (2012): 1080-1087.

References

Noonan, Curtis W., Erin O. Semmens, Paul Smith, Solomon W. Harrar, Luke Montrose, Emily Weiler, Marcy McNamara, and Tony J. Ward. "Randomized trial of interventions to improve childhood asthma in homes with wood-burning stoves." Environmental health perspectives 125, no. 9 (2017): 097010. ([PubMed](https://pubmed.ncbi.nlm.nih.gov/28935614/));

Examples

```
data(ARTIS)
head(ARTIS)
```

ConfInterval *Confidence Interval*

Description

Construct confidence intervals for effect sizes.

Usage

```
ConfInterval(object, level, side="two.sided",
adjust=NULL)
```

Arguments

object	a fitted model object from ncda().
level	the confidence level required.
side	a character string specifying the side of the confidence bound, must be one of "two.sided" (default), "left" or "right".
adjust	an optional character string specifying the multiple adjustment method, by default there is no adjustment, if specified, must be one of "Bonferroni" or "Working-Hotelling". You can specify just the initial letter.

Value

A list or a vector. If the confidence interval is two-sided, lower and upper bounds are stored in lists for each nonparametric effect size estimate. Otherwise, the lower/upper bounds are stored in vectors in the order of the effect size estimates.

Examples

```
skin_analysis <- ncda(score~tx, skin, intervention, subject,
                      indicator=c("control","treatment"),
                      Contrast=matrix(c(1,-1), nrow = 1))
ConfInterval(skin_analysis,0.95)

ARTIS_analysis <- ncda(emot_pqol~tx, ARTIS, intervention, homeid,
                       indicator = c("0","1"))
ConfInterval(ARTIS_analysis,0.95)
ConfInterval(ARTIS_analysis,0.95,"two.sided","B")
ConfInterval(ARTIS_analysis,0.95,"left","W")
```

ncda *Nonparametric Clustered Data Analysis*

Description

Main function to calculate nonparametric effect sizes and conduct hypothesis tests.

Usage

```
ncda(formula,data,period,subject,indicator=NULL,Contrast=NULL)
```

Arguments

formula	An object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given under 'Details'.
data	a data frame in the long format.
period	time indicator variable.
subject	subject or cluster ID
indicator	an optional vector of characters indicating the order of pre and post intervention period; must match the levels of period argument if specified; if not specified, the pre and post intervention period will be ordered in the alphabet order by default
Contrast	an optional contrast matrix for effect sizes.

Details

The model has the form response ~ tx where response is the (numeric) response variable and tx is the treatment variable.

Value

An object with effect sizes and other test details.

References

Cui, Yue, Frank Konietschke, and Solomon W. Harrar. "The nonparametric Behrens–Fisher problem in partially complete clustered data." Biometrical Journal 63.1 (2021): 148-167.

Harrar, Solomon W., and Yue Cui. "Nonparametric methods for clustered data in pre-post intervention design." Journal of Statistical Planning and Inference 222 (2023): 1-21.

Examples

```
ARTIS_analysis <- ncda(symptoms_pqol~tx, ARTIS, intervention, homeid,
                       indicator=c("0","1"),
                       Contrast=matrix(c(1,-1,1,-1,1,-1), nrow = 1))
names(ARTIS_analysis)
ARTIS_analysis$p.vector

skin_analysis <- ncda(score~tx, skin, intervention, subject,
                      indicator=c("control","treatment"),
                      Contrast=matrix(c(1,-1), nrow = 1))
skin_analysis$TotalSampleSize
skin_analysis$p.vector
```

Plot.box *Box plots.*

Description

Box plot of the input data set by treatments and time period.

Usage

```
Plot.box(object)
```

Arguments

object	a fitted model object from ncda() or a processed data set from ProcessData()

Value

Box plots.

Examples

```
#Plot from analysis object

ARTIS_analysis <- ncda(emot_pqol~tx, ARTIS, intervention, homeid,
                       indicator = c("0","1"))
Plot.box(ARTIS_analysis)

skin_analysis <- ncda(score~tx, skin, intervention, subject,
                      indicator = c("control","treatment"))
Plot.box(skin_analysis)
# Plot from processed data set
ARTIS_result <- ProcessData(ARTIS, tx, intervention, homeid, symptoms_pqol,
                            indicator = c("0","1"))
skin_result <- ProcessData(skin, tx, intervention, subject, score,
                           indicator = c("control","treatment"))
Plot.box(ARTIS_result)
Plot.box(skin_result)
```

`Plot.ConfInt` *Bar plots for two-sided confidence intervals*

Description

Bar plots for two-sided confidence intervals

Usage

```
Plot.ConfInt(object, level, side="two.sided",
adjust=NULL)
```

Arguments

`object`	a fitted model object from ncda().
`level`	the confidence level required.
`side`	a character string specifying the side of the confidence bound, must be one of "two.sided" (default), "left" or "right".
`adjust`	an optional character string specifying the multiple adjustment method, by default there is no adjustment, if specified, must be one of "Bonferroni" or "Working-Hotelling". You can specify just the initial letter.

Value

Bar plots.

Examples

```
skin_analysis <- ncda(score~tx, skin, intervention, subject,
                      indicator=c("control","treatment"),
                      Contrast=matrix(c(1,-1), nrow = 1))
Plot.ConfInt(skin_analysis,0.95,"Two-Sided")

ARTIS_analysis <- ncda(emot_pqol~tx, ARTIS, intervention, homeid,
                       indicator = c("0","1"))
Plot.ConfInt(ARTIS_analysis,0.95,"Two-Sided")
Plot.ConfInt(ARTIS_analysis,0.95,"Two-Sided","Bonferroni")
Plot.ConfInt(ARTIS_analysis,0.95,"Two-Sided","Working-Hotelling")
```

`ProcessData` *Process data set.*

Description

Sample size and cluster size calculation for the imported data set.

Usage

```
ProcessData(data, tx, period, subject, resp, indicator=NULL)
```

Arguments

`data`	a data frame in the long format.
`tx`	treatment variable.
`period`	time indicator variable.
`subject`	subject or cluster ID
`resp`	response variable to be analyzed.
`indicator`	an optional vector of characters indicating the order of pre and post intervention period; must match the levels of period argument if specified; if not specified, the pre and post intervention period will be ordered in the alphabet order by default

Value

a list containing the following components:

`trt` number of treatments

`nc` complete cluster sample size within each treatment group

`n1` incomplete cluster sample size pre intervention within each treatment group

`n2` incomplete cluster sample size post intervention within each treatment group

`m1c` complete cluster sizes pre-intervention within each treatment group

`m2c` complete cluster size post-intervention within each treatment group

`m1i` incomplete cluster sizes pre-intervention within each treatment group

`m2i` incomplete cluster sizes post-intervention within each treatment group

`x1c` complete data pre-intervention within each treatment group

`x2c` complete data post-intervention within each treatment group

`x1i` incomplete data pre-intervention within each treatment group

`x2i` incomplete data post-intervention within each treatment group

Examples

```
ARTIS_result <- ProcessData(ARTIS, tx, intervention, homeid, symptoms_pqol,
                            c("0","1"))
names(ARTIS_result)
skin_result <- ProcessData(skin, tx, intervention, subject, score,
                           c("control","treatment"))
skin_result$nc
skin_result$n1
skin_result$n2
```

---

skin *Skin Irritation Data*

Description

The data set is a re-simulated part from an ongoing neuodermatitis study where researchers investigate the efficacy of an ointment in reducing the severity of skin irritation on the backs of the hands of 25 neurodermatitis patients, where 10 patients' backs of the hands were rubbed with the ointment and 15 were not. The response is a BI-RADS rating score and the lower the score the better the clinical record. Every remarkable skin irritation was graded on every patients back of the hands and thus, the numbers of replicates differ across the patients.

Usage

```
skin
```

Format

A data frame with 107 rows and 4 variables:

`tx` treatment group

`intervention` intervention period indicator

`subject` subject ID

`score` BI-RADS rating score, where 1 = very mild irritation, 2 =slight irritation, 3 =mild irritation, 4 =heavy irritation and 5 =severe irritation

References

Roy, A, Harrar, SW, Konietschke, F. The nonparametric Behrens-Fisher problem with dependent replicates. Statistics in Medicine. 2019; 38: 4939– 4962. https://doi.org/10.1002/sim.8343

Examples

```
data(skin)
head(skin)
```

Index