Package ‘permuco’

December 19, 2019

Type Package

Title Permutation Tests for Regression, (Repeated Measures) ANOVA/ANCOVA and Comparison of Signals

Version 1.1.0

Date 2019-12-19


License GPL (>= 2)

Imports permute,Matrix,stats,graphics

LazyData TRUE

RoxygenNote 7.0.2

Suggests R.rsp

VignetteBuilder R.rsp

NeedsCompilation no

Author Jaromil Frossard [aut, cre], Olivier Renaud [aut]

Maintainer Jaromil Frossard <jaromil.frossard@unige.ch>

Repository CRAN

Date/Publication 2019-12-19 13:40:05 UTC
**R topics documented:**

aovperm ................................................................. 2
as.Pmat ................................................................. 4
attentionshifting_design ............................................. 5
attentionshifting_signal ............................................ 6
clusterlm ................................................................. 6
compute_clustermass ................................................ 9
compute_minP .......................................................... 10
compute_tfce .......................................................... 10
compute_troendle ..................................................... 11
emergencycost ........................................................ 11
jpah2016 ................................................................. 12
lmperm ................................................................. 13
plot.clusterlm ......................................................... 14
plot.lmperm ............................................................. 16
Pmat ................................................................. 16
print.clusterlm ........................................................ 17
summary.clusterlm ................................................... 18

---

**Description**

Provides p-values for omnibus tests based on permutations for factorial and repeated measures ANOVA. This function produces the F statistics, parametric p-values (based on Gaussian and sphericity assumptions) and p-values based on the permutation methods that handle nuisance variables.

**Usage**

```r
aovperm(formula, data = NULL, np = 5000, method = NULL, ...)
```

**Arguments**

- `formula`: A formula object. The formula for repeated measures ANOVA should be written using the same notation as `aov` by adding `+Error(id/within)`, where `id` is the factor that identify the subjects and `within` is the within factors.
- `data`: A data frame or matrix.
- `np`: The number of permutations. Default value is 5000.
- `method`: A character string indicating the method used to handle nuisance variables. Default is NULL and will change if to "freedman_lane" for the fixed effects model and "Rd_kheradPajouh_renaud" for the random effect models. See Details for other methods.
- `...`: Further arguments, see details.
Details

The following methods are available for the fixed effects model defined as \( y = D \eta + X \beta + \epsilon \). If we want to test \( \beta = 0 \) and take into account the effects of the nuisance variables \( D \), we transform the data:

<table>
<thead>
<tr>
<th>method argument</th>
<th>y*</th>
<th>D*</th>
<th>X*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;draper_stoneman&quot;</td>
<td>y</td>
<td>D</td>
<td>PX</td>
</tr>
<tr>
<td>&quot;freedman_lane&quot;</td>
<td>(( H_D + PR_D ))y</td>
<td>D</td>
<td>X</td>
</tr>
<tr>
<td>&quot;manly&quot;</td>
<td>Py</td>
<td>D</td>
<td>X</td>
</tr>
<tr>
<td>&quot;terBraak&quot;</td>
<td>(( H_{X,D} + PR_{X,D} ))y</td>
<td>D</td>
<td>X</td>
</tr>
<tr>
<td>&quot;kennedy&quot;</td>
<td>PR_{Dy}</td>
<td>R_{D}</td>
<td>X</td>
</tr>
<tr>
<td>&quot;huh_jhun&quot;</td>
<td>PV'R_{Dy}</td>
<td>V'R_{D}</td>
<td>X</td>
</tr>
<tr>
<td>&quot;dekker&quot;</td>
<td>y</td>
<td>D</td>
<td>PR_{DX}</td>
</tr>
</tbody>
</table>

The following methods are available for the random effects model \( y = D \eta + X \beta + E \kappa + Z \gamma + \epsilon \). If we want to test \( \beta = 0 \) and take into account the effect of the nuisance variable \( D \) we can transform the data by permutation:

<table>
<thead>
<tr>
<th>method argument</th>
<th>y*</th>
<th>D*</th>
<th>X*</th>
<th>E*</th>
<th>Z*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Rd_kheradPajouh_renaud&quot;</td>
<td>PR_{Dy}</td>
<td>R_{D}</td>
<td>X</td>
<td>R_{D}</td>
<td>R_{D}</td>
</tr>
<tr>
<td>&quot;Rde_kheradPajouh_renaud&quot;</td>
<td>PR_{D,Ey}</td>
<td>R_{D,E}</td>
<td>X</td>
<td>R_{D,E}</td>
<td>R_{D,E}</td>
</tr>
</tbody>
</table>

Other arguments could be pass in ...:

- \( P \): a matrix, of class matrix or Pmat, containing the permutations (for the reproductibility of the results). The first column must be the identity permutation (not checked). \( P \) overwrites np argument.
- rnd_rotation : a random matrix of size \( n \times n \) to compute the rotation used for the "huh_jhun" method.
- coding_sum : a logical set to TRUE defining the coding of the design matrix to contr.sum to test the main effects. If it is set to FALSE the design matrix is computed with the coding defined in the dataframe. The tests of simple effects are possible with a coding of the factors of the dataframe set to contr.treatment.

Value

A lmperm object containing most of the objects given in an lm object, an ANOVA table with parametric and permutation p-values, the test statistics and the permutation distributions.

Author(s)

jaromil.frossard@unige.ch

See Also

lmperm, plot.lmperm
Examples

```r
## data
data("emergencycost")

## centrering the covariate to the mean
evacuity$LOSc <- scale(evacuity$LOS, scale = FALSE)

## ANCOVA
## Warning : np argument must be greater (recommendation: np>=5000)
mod_cost_0 <- aovperm(cost ~ LOSc*sex*insurance, data = evacuity, np = 2000)
mod_cost_0

## Testing at 14 days
evacuity$LOS14 <- evacuity$LOS - 14

mod_cost_14 <- aovperm(cost ~ LOS14*sex*insurance, data = evacuity, np = 2000)
mod_cost_14

## Effect of sex within the public insured
contrasts(evacuity$insurance) <- contr.treatment
contrasts(evacuity$sex) <- contr.sum
vacuity$insurance <- relevel(evacuity$insurance, ref = "public")

mod_cost_se <- aovperm(cost ~ LOSc*sex*insurance, data = evacuity, np = 2000, coding_sum = FALSE)
mod_cost_se

## Repeated measures ANCOVA
## data
data(jpah2016)

## centrering the covariate
jpah2016$bmic <- scale(jpah2016$bmi, scale = FALSE)

## Warning : np argument must be greater (recommendation: np>=5000)
mod_jpah2016 <- aovperm(iapa ~ bmic*condition*time+ Error(id/(time)),
                        data = jpah2016, method = "Rd_kheradPajouh_renaud")
mod_jpah2016
```

---

**as.Pmat**

*Method to convert into Pmat object.*

**Description**

Convert a matrix into a Pmat object.
**attentionshifting_design**

**Usage**

\[
\text{as.Pmat}(x)
\]

**Arguments**

- \(x\) a matrix.

---

**attentionshifting_design**

_Dataset of the design for the data attentionshifting_signal_

**Description**

Design of an experiment measuring the EEG brain activity of 15 participants who have been shown images of neutral and angry faces. Those faces were shown at a different visibility 16ms and 166ms and were displayed either to the left or to the right of a screen. The laterality, sex, age, and 2 measures of anxiety for each subject are also available. The amplitude of the EEG recording are located in the dataset `attentionshifting_signal`.

- id : identifier of the subject.
- visibility : time of exposure to the image (16ms: subliminal or 166ms: supraliminal).
- emotion : type of emotion of the image (angry or neutral).
- direction : position of image on the screen (left or right).
- laterality_id : measure of laterality of the subject.
- age : age of the subject.
- sex : sex of the subject.
- STAIS_state : measure of the state of anxiety of the subject.
- STAIS_trait : measure of the personality trait of anxiety of the subject.

**Usage**

\[
data(\text{attentionshifting_design})
\]

**Format**

A data frame with 120 rows and 10 variables.
attentionshifting_signal

Dataset containing the event-related potential of the electrod O1 of a control experiment.

Description

The ERP of the electrod O1 of an experiment in attention shifting. This dataset contains the amplitude of the signals sampled at 1024 Hz. The design of the experiment is given in the dataset attentionshifting_design.

Usage

data(attentionshifting_signal)

Format

A data frame with 120 rows and 819 variables.

Details

- ERP (in μV) of the electrod O1 measured from -200 to 600 timeframes before and after the onset of the stimulus.

clusterlm

Cluster test for longitudinal data

Description

Compute the cluster-mass test for longitudinal linear model.

Usage

clusterlm(
  formula,
  data = NULL,
  np = 5000,
  method = NULL,
  test = "fisher",
  threshold = NULL,
  aggr_FUN = NULL,
  multcomp = "clustermass",
  ...)
)
Arguments

formula A formula object where the left part is a matrix defined in the global environment.
data A data frame for the independent variables.
np The number of permutations. Default value is 5000.
method A character string indicating the method used to handle nuisance variables. Default is NULL and will switch to "freedman_lane" for the fixed effects model and to "Rd_kheradPajouh_renaud" for the repeated measures ANOVA. See lmperm or aovperm for details on the permutation methods.
test A character string to specify the name of the test. Default is "fisher". "t" is available for the fixed effects model.
threshold A numerical value that specify the threshold for the "clustermass" multiple comparisons procedure. If it is a vector each value will be associated to an effect. If it is scalar the same threshold will be used for each test. Default value is NULL and will compute a threshold based on the 0.95 quantile of the chosen test statistic.
aggr_FUN A function used as mass function. It should aggregate the statistics of a cluster into one scalar. Default is the sum of squares for t statistic and sum for F statistic.
multcomp A vector of character defining the methods of multiple comparisons to compute. Default is "clustermass", and the additional options are available: "tfce", "bonferroni", "holm", "troendle", "minP" and "benjamini_hochberg".

Details

The random effects model is only available with a F statistic.

Other arguments could be pass in ... :

P : A matrix containing the permutation of class matrix or Pmat; which is used for the reproducibility of the results. The first column must be the identity. P overwrites np argument.

rnd_rotation : A matrix of random value to compute a rotation of size $n \times n$ that will be used for the "huh_jhun" method.

p_scale = FALSE : if set to TRUE, the several multiple comparisons procedures are compute on the 1 -p scale, where p is the p-value. The threshold have to be set between 0 and 1 (eg: threshold = 0.95). The function aggr_FUN should be big when there is evidence against the null (eg: aggr_FUN = function(p)sum(abs(log(1-p)))). Moreover under the probability scale the cluster mass statistics is sensitive to the number permutations.

H, E, ndh : the parameters used for the "tfce" method. Default values are set to $H = 2$ for the height parameter, to $E = 0.5$ for the extend parameter and to $ndh = 500$ for the number terms to approximate the integral.
alpha = 0.05: the type I error rate. Used for the troendle multiple comparisons procedure.

return_distribution = FALSE: return the permutation distribution of the statistics. Warnings: return one high dimensional matrices (number of test times number of permutation) for each test. coding_sum: a logical defining the coding of the design matrix to contr.sum: set by default to TRUE for ANOVA (when the argument test is "fisher") to tests main effects and is set to FALSE when test is "t". If coding_sum is set to FALSE the design matrix is computed with the coding defined in the dataframe and the tests of simple effects are possible with a coding of the dataframe set to contr.treatment.

Value

A clusterlm object. Use the plot.clusterlm or summary.clusterlm method to see results of the tests.

Author(s)

jaromil.frossard@unige.ch

References


See Also

plot.clusterlm summary.clusterlm

Examples

## Cluster-mass for repeated measures ANOVA
## Warning: np argument must be greater (recommendation: np >= 5000)
electrod_O1 <- clusterlm(attentionshifting_signal ~ visibility*emotion*direction + Error(id/(visibility*emotion*direction)), data = attentionshifting_design, np = 50)

## Results
plot(electrod_O1)

## Results with labels on the x axis that represent seconds from time-locked event:
plot(electrod_O1, nbbaselinepts = 200, nbptsperunit = 1024)

## Tables of clusters
electrod_O1

## Not run:
## Change the function of the aggregation
compute_clustermass

## Sum of squares of F statistics

electrod_O1_sum <- clusterlm(attentionshifting_signal ~ visibility*emotion*direction + Error(id/(visibility*emotion*direction)), data = attentionshifting_design, aggr_FUN = function(x)sum(x^2))

## Length of the cluster

electrod_O1_length <- clusterlm(attentionshifting_signal ~ visibility*emotion*direction + Error(id/(visibility*emotion*direction)), data = attentionshifting_design, aggr_FUN = function(x)length(x))

## All multiple comparisons procedures for repeated measures ANOVA
## Permutation method "Rde_kheradPajouh_renaud"

full_electrod_O1 <- clusterlm(attentionshifting_signal ~ visibility*emotion*direction + Error(id/(visibility*emotion*direction)), data = attentionshifting_design, method = "Rde_kheradPajouh_renaud", multcomp = c("troendle", "tfce", "clustermass", "bonferroni", "holm", "benjamini_hochberg"))

## End(Not run)

---

### compute_clustermass

#### Clustermass test correction

**Description**

Compute the clustermass test correction given a matrix a permuted statistical signals.

**Usage**

```r
compute_clustermass(distribution, threshold, aggr Fun, alternative = "greater")
```

**Arguments**

- `distribution`: A matrix of permuted statistical signal. The first row indicating the observed statistics.
- `threshold`: A scalar that represents the threshold to create the clusters.
- `aggr Fun`: A function to compute the clustermasses. See details for examples.
- `alternative`: A character string indicating the alternative hypothesis. Default is "greater". Choose between "greater", "less" or "two.sided".

**Details**

The `aggr Fun` argument may take predefined function as the sum: `aggr FUN = sum` and also user-defined function as the sum of squares: `aggr FUN = function(x)sum(x^2)`
**compute_minP**  
*The min-P correction*

**Description**
Compute the min-P correction given a matrix a permuted statistics.

**Usage**
\[ \text{compute\_minP(distribution, alternative)} \]

**Arguments**

- **distribution**  A matrix of permuted statistical signal. The first row indicating the observed statistics.
- **alternative**  A character string indicating the alternative hypothesis. Default is "greater". Choose between "greater", "less" or "two.sided".

**compute_tfce**  
*Threshold-Free Cluster-Enhancement correction*

**Description**
Compute the TFCE correction given a matrix a permuted statistical signals.

**Usage**
\[ \text{compute\_tfce(distribution, E, H, ndh)} \]

**Arguments**

- **distribution**  A matrix of permuted statistical signal. The first row indicating the observed statistics.
- **E**  A scalar that represent the extend parameter of the TFCE transformation. Default is \( E = 0.5 \).
- **H**  A scalar that represent the height parameter of the TFCE transformation. Default is \( H = 1 \).
- **ndh**  The number of terms in the approximation of the integral.
**compute_troendle**

*The Troendle’s correction*

**Description**

Compute the Troendle’s correction given a matrix a permuted statistics.

**Usage**

```r
compute_troendle(distribution, alternative)
```

**Arguments**

- `distribution`: A matrix of permuted statistical signal. The first row indicating the observed statistics.
- `alternative`: A character string indicating the alternative hypothesis. Default is "greater". Choose between "greater", "less" or "two.sided".

**emergencycost**

*Dataset of cost of emergency patients.*

**Description**

Observational data from 176 emergency patients with variables:

**Usage**

```r
data(emergencycost)
```

**Format**

A data frame with 176 rows and 5 variables.

**Details**

- `sex`.
- `age`.
- `insurance`: the type of insurance, private or semi private (`semi_private`) or public (`public`).
- `LOS`: the length of the stay in days.
- `cost`: the cost in CHF.

**References**

Dataset of a control study in psychology.

Description
A subset of a control experiment measuring the impulsive approach tendencies toward physical activity or sedentary behaviors.

Usage
data(jpah2016)

Format
A data frame with 38 rows and 8 variables.

Details
- id identifier of the subject.
- bmi body mass index.
- age.
- sex.
- condition the experimental condition where the task was to approach physical activity and avoid sedentary behavior (ApSB_AvPA), approach sedentarity behavior and avoid physical activity (ApPA_AvSB), and a control condition (control).
- time pre, post.
- iapa measure of impulsive approach tendencies toward physical activity (dependant variable).
- iasb measure of impulsive approach tendencies toward sedentary behavior (dependant variable).

References
Imperm

Permutation tests for regression parameters

Description

Compute permutation marginal tests for linear models. This function produces t statistics with univariate and bivariate p-values. It gives the choice between multiple methods to handle nuisance variables.

Usage

Imperm(formula, data = NULL, np = 5000, method = NULL, ...)

Arguments

- formula: A formula object.
- data: A data frame or matrix.
- np: The number of permutations. Default value is 5000.
- method: A character string indicating the method used to handle nuisance variables. Default is "freedman_lane". For the other methods, see details.
- ...: Further arguments, see details.

Details

The following methods are available for the fixed effects model defined as \( y = D \eta + X \beta + \epsilon \). If we want to test \( \beta = 0 \) and take into account the effects of the nuisance variables \( D \), we transform the data:

<table>
<thead>
<tr>
<th>method argument</th>
<th>( y^* )</th>
<th>( D^* )</th>
<th>( X^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;draper_stoneman&quot;</td>
<td>( y )</td>
<td>( D )</td>
<td>( PX )</td>
</tr>
<tr>
<td>&quot;freedman_lane&quot;</td>
<td>( (H_D + PR_D)y )</td>
<td>( D )</td>
<td>( X )</td>
</tr>
<tr>
<td>&quot;manly&quot;</td>
<td>( Py )</td>
<td>( D )</td>
<td>( X )</td>
</tr>
<tr>
<td>&quot;terBraak&quot;</td>
<td>( (H_{X,D} + PR_{X,D})y )</td>
<td>( D )</td>
<td>( X )</td>
</tr>
<tr>
<td>&quot;kennedy&quot;</td>
<td>( PR_{DY} )</td>
<td>( R_{D}X )</td>
<td></td>
</tr>
<tr>
<td>&quot;huh_jhun&quot;</td>
<td>( PV'R_{DY} )</td>
<td>( V'R_{D}X )</td>
<td></td>
</tr>
<tr>
<td>&quot;dekker&quot;</td>
<td>( y )</td>
<td>( D )</td>
<td>( PR_{D}X )</td>
</tr>
</tbody>
</table>

Other arguments could be passed in ...:

- \( P \): a matrix containing the permutations of class matrix or Pmat for the reproducibility of the results. The first column must be the identity. \( P \) overwrites np argument.

- \( rnd_rotation \): a random matrix of size \( n \times n \) to compute the rotation used for the "huh_jhun" method.
Plot cluster or parameters.

Description

Plot method for class clusterlm.
## Usage

### S3 method for class 'clusterlm'

```r
plot(
  x,
  effect = "all",
  type = "statistic",
  multcomp = x$multcomp[1],
  alternative = "two.sided",
  enhanced_stat = FALSE,
  nbbaselinepts = 0,
  nbptsperunit = 1,
  distinctDVs = NULL,
  ...
)
```

### Arguments

- **x**: A `clusterlm` object.
- **effect**: A vector of character naming the effects to display. Default is "all".
- **type**: A character string that specified the values to highlight. "statistic" or "coef" are available. Default is "statistic".
- **multcomp**: A character string specifying the method use to correct the p-value. It should match the one computed in the object. Default is the (first) method in the call to `clusterlm`. See `clusterlm`.
- **alternative**: A character string specifying the alternative hypothesis for the t-test. The available options are "greater", "less" and "two.sided". Default is "two.sided".
- **enhanced_stat**: A logical. Default is `FALSE`. If `TRUE`, the enhanced statistic will be plotted otherwise it will plot the observed statistic. Change for the "tfce" or the "clustermass" multiple comparisons procedures.
- **nbbaselinepts**: An integer. Default is 0. If the origin of the x axis should be shifted to show the start of the time lock, provide the number of baseline time points.
- **nbptsperunit**: An integer. Default is 1. Modify this value to change the scale of the label from the number of points to the desired unit. If points are e.g. sampled at 1024Hz, set to 1024 to scale into seconds and to 1.024 to scale into milliseconds.
- **distinctDVs**: Boolean. Should the DVs be plotted distinctively, i.e. should the points be unlinked and should the name of the DVs be printed on the x axis? Default is `FALSE` if the number of DV is large than 15 or if the method is "clustermass" or "tfce".
- **...**: further argument pass to plot.
Description
Show the density of statistics and the test statistic.

Usage
```r
## S3 method for class 'lmperm'
plot(x, FUN = density, ...)
```

Arguments
- `x`: A "lmperm" object.
- `FUN`: A function to compute the density. Default is `density`.
- `...`: Further arguments pass to `plot`.

Details
Other argument can be pass to the function:
- `effect`: a vector of character string indicating the name of the effect to plot.

---

Pmat
Create a set of permutations.

Description
Compute a permutation matrix used as argument in `aovperm`, `Imperm`, `clusterlm` functions. The first column represents the identity permutation.

Usage
```r
Pmat(np = 5000, n, type = "default")
```

Arguments
- `np`: A numeric value for the number of permutations. Default is 5000.
- `n`: A numeric value for the number of observations.
- `type`: A character string to specify the type of matrix. See Details.
Details

type can set to:
"default" : np random with replacement permutations among the n! permutations.
"all" : all n! possible permutations.

Value

A matrix n x np containing the permutations/coinflips. First permutation is the identity.

Examples

```r
## data
data("emergencycost")

## Create a set of 2000 permutations
set.seed(42)
pmat = Pmat(np = 2000, n = nrow(emergencycost))

## centring the covariate to the mean
emergencycost$LOSc <- scale(emergencycost$LOS, scale = FALSE)

## ANCOVA
mod_cost_0 <- aovperm(cost ~ LOSc*sex*insurance, data = emergencycost, np = 2000)
mod_cost_1 <- aovperm(cost ~ LOSc*sex*insurance, data = emergencycost, P = pmat)
mod_cost_2 <- aovperm(cost ~ LOSc*sex*insurance, data = emergencycost, P = pmat)

## Same p-values for both models 1 and 2 but different of model 0
mod_cost_0
mod_cost_1
mod_cost_2
```

print.clusterlm

Print clusterlm object.

Description

Display the corrected p-values for each effects. Results of the "clustermass" procedure.

Usage

```r
## S3 method for class 'clusterlm'
print(x, multcomp = NULL, alternative = "two.sided", ...)
```
**summary.clusterlm**

**Arguments**

- **x**: A `clusterlm` object.
- **multcomp**: A character string indicating the multiple comparison procedure to print. Default is NULL, a print the first multiple comparisons procedure of the `clusterlm` object.
- **alternative**: A character string indicating the alternative hypothesis. Choose between "two.sided", "greater", "less". Default is "two.sided".
- **...**: Further arguments pass to `print`.

**Description**

Display the corrected p-values for each effects.

**Usage**

```r
## S3 method for class 'clusterlm'
summary(
  object,
  alternative = "two.sided",
  multcomp = NULL,
  table_type = NULL,
  ...
)
```

**Arguments**

- **object**: A `clusterlm` object.
- **alternative**: A character string indicating the alternative hypothesis. Choose between "two.sided", "greater", "less". Default is "two.sided".
- **multcomp**: A character string indicating the multiple comparison procedure to display.
- **table_type**: A character string indicating the type of table to display. Choose between "cluster", which aggregates test into pseudo-clusters (see details for the interpretations) or "full" which displays the p-values for all tests. See details for default values.
- **...**: Further arguments see details.

**Details**

It creates the full table when the number of tests is \( \leq 15 \) and creates a table of pseudo-clusters otherwise. Note that for the "troendle" method is not based on clustering of the data and the table of pseudo-clusters should only be used to facilitate the reading of the results.
Value

A table for each effect indicating the statistics and p-values of the clusters.
Index

aov, 2
aovperm, 2, 7, 14, 16
as.Pmat, 4
attentionshifting_design, 5, 6
attentionshifting_signal, 5, 6

cluster1m, 6, 15, 16
compute_clustermass, 9
compute_minP, 10
compute_tfce, 10
compute_troendle, 11

density, 16

emergencycost, 11

jpah2016, 12

lm, 3
lmperm, 3, 7, 13, 16

plot.clusterlm, 8, 14
plot.lmperm, 3, 14, 16
Pmat, 16
print.clusterlm, 17

summary.clusterlm, 8, 18