# Package ‘MANOVA.RM’

August 28, 2019

<table>
<thead>
<tr>
<th>Type</th>
<th>Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Resampling-Based Analysis of Multivariate Data and Repeated Measures Designs</td>
</tr>
<tr>
<td>Version</td>
<td>0.3.4</td>
</tr>
<tr>
<td>Date</td>
<td>2019-08-27</td>
</tr>
<tr>
<td>Author</td>
<td>Sarah Friedrich, Frank Konietschke, Markus Pauly</td>
</tr>
<tr>
<td>Maintainer</td>
<td>Sarah Friedrich <a href="mailto:sarah.friedrich@med.uni-goettingen.de">sarah.friedrich@med.uni-goettingen.de</a></td>
</tr>
<tr>
<td>Depends</td>
<td>R (&gt;= 3.4.0)</td>
</tr>
</tbody>
</table>

**Description**

Implemented are various tests for semi-parametric repeated measures and general MANOVA designs that do neither assume multivariate normality nor covariance homogeneity, i.e., the procedures are applicable for a wide range of general multivariate factorial designs. In addition to asymptotic inference methods, novel bootstrap and permutation approaches are implemented as well. These provide more accurate results in case of small to moderate sample sizes. Furthermore, post-hoc comparisons are provided for the multivariate analyses.


**License**

GPL-2 | GPL-3

**Imports**

plyr (>= 1.8.3), MASS (>= 7.3-43), Matrix (>= 1.2-2), magic (>= 1.5-6), plotrix (>= 3.5-12), parallel, methods, ellipse, multcomp, data.table

**LazyData**

TRUE

**Suggests**

RGtk2 (>= 2.20.31), knitr, rmarkdown, HSAUR, tidyr, GFD, testthat, tidyverse

**Encoding**

UTF-8

**RoxygenNote**

6.1.1

**VignetteBuilder**

knitr, rmarkdown

**URL**

http://github.com/smn74/MANOVA.RM

**BugReports**

http://github.com/smn74/MANOVA.RM/issues

**NeedsCompilation**

no
The `conf.reg()` function calculates confidence regions for contrasts in multivariate factorial designs. In the two-dimensional case, confidence ellipsoids can be plotted via the generic `plot()` function.

**Description**

The `conf.reg()` function calculates confidence regions for contrasts in multivariate factorial designs. In the two-dimensional case, confidence ellipsoids can be plotted via the generic `plot()` function.

**Usage**

```r
conf.reg(object, nullhypo)
```

**Arguments**

- **object**: A MANOVA object.
- **nullhypo**: In scenarios with more than one factor, the null hypothesis, i.e., the contrast of interest must be specified.
**Value**

A `confreg` object containing the following components:

- **center**: The center of the confidence ellipsoid.
- **scale**: The scaling factors for the axis of the confidence ellipsoid calculated as $\sqrt{\lambda \ast c/N}$, where $\lambda$ are the eigenvalues, $c$ denotes the bootstrap quantile and $N$ is the total sample size. See Friedrich and Pauly (2018) for details.
- **eigenvectors**: The corresponding eigenvectors, which determine the axes of the ellipsoid.

**References**


**Examples**

```r
data(EEG)
EEG_mod <- MANOVA(resp ~ sex * diagnosis,
                  data = EEG, subject = "id", resampling = "paramBS",
                  alpha = 0.05, iter = 100, CPU = 1)
conf.reg(EEG_mod, nullhypo = "sex")
```

---

**EEG Measurements in Patients with Alzheimer’s Disease (long format)**

**Description**

At the Department of Neurology, University Clinic of Salzburg, 160 patients were diagnosed with either AD, MCI, or SCC, based on neuropsychological diagnostics. This data set contains z-scores for brain rate and Hjorth complexity, each measured at frontal, temporal and central electrode positions and averaged across hemispheres. In addition to standardization, complexity values were multiplied by -1 in order to make them more easily comparable to brain rate values: For brain rate we know that the values decrease with age and pathology, while Hjorth complexity values are known to increase with age and pathology. The three between-subjects factors considered were sex (men vs. women), diagnosis (AD vs. MCI vs. SCC), and age (< 70 vs. >= 70 years). Additionally, the within-subjects factors region (frontal, temporal, central) and feature (brain rate, complexity) structure the response vector.

**Usage**

```r
data(EEG)
```
Format

A data frame with 960 rows and 7 variables:

- resp  EEG measurements
- sex   sex of the patient
- age   age of the patient, coded as 0 for less than 70 years and 1 for \( \geq 70 \) years
- diagnosis  neuropsychological diagnosis, AD for Alzheimer’s Disease, MCI for mild cognitive impairment or SCC for subjective cognitive complaints without clinically significant deficits
- region  brain region of the EEG measurements, one of "temporal", "frontal" and "central"
- feature  feature of the EEG measurements, either "brainrate" or "complexity"
- id     Subject id

Source


Examples

```r
library(tidyverse)
ggplot(EEG, aes(x=sex, y=resp)) + geom_point(alpha=0.5) + facet_grid(region+feature~diagnosis) +
  stat_summary(fun.y = mean, fun.ymin = min, fun.ymax = max, colour = "red")
```

Usage

data(EEGwide)

Description

At the Department of Neurology, University Clinic of Salzburg, 160 patients were diagnosed with either AD, MCI, or SCC, based on neuropsychological diagnostics. This data set contains z-scores for brain rate and Hjorth complexity, each measured at frontal, temporal and central electrode positions and averaged across hemispheres. In addition to standardization, complexity values were multiplied by -1 in order to make them more easily comparable to brain rate values: For brain rate we know that the values decrease with age and pathology, while Hjorth complexity values are known to increase with age and pathology. The three between-subjects factors considered were sex (men vs. women), diagnosis (AD vs. MCI vs. SCC), and age (\(< 70\) vs. \(\geq 70\) years). Additionally, the within-subjects factors region (frontal, temporal, central) and feature (brain rate, complexity) structure the response vector.
Format

A data frame with 160 rows and 9 variables:

- `brainrate_temporal` EEG measurements for brainrate in temporal regions
- `brainrate_frontal` EEG measurements for brainrate in frontal regions
- `brainrate_central` EEG measurements for brainrate in central regions
- `complexity_temporal` EEG measurements for complexity in temporal regions
- `complexity_frontal` EEG measurements for complexity in frontal regions
- `complexity_central` EEG measurements for complexity in central regions
- `sex` sex of the patient
- `age` age of the patient
- `diagnosis` neuropsychological diagnosis, AD for Alzheimer’s Disease, MCI for mild cognitive impairment or SCC for subjective cognitive complaints without clinically significant deficits
- `AgeGroup` categorized age, coded as 0 for less than 70 years and 1 for >= 70 years

Details

Note that this data set contains exactly the same data as the data set 'EEG', only the format is different. The transformation between the different formats can be achieved using, e.g., the `tidyverse` package.

Source


@examples

```r
library("ggplot2")
qplot(data = EEGwide, diagnosis)
```

GUI.MANOVA

A graphical user interface for the MANOVA() function

Description

This function provides a graphical user interface for calculating statistical tests for multivariate data.

Usage

`GUI.MANOVA()`

Details

The function produces a GUI for the calculation of the test statistics. Data can be loaded via the "load data" button. The formula, number of resampling iterations (default: 10,000) and the significance level alpha (default: 0.05) need to be specified. Furthermore, the column name specifying the subjects in the data needs to be provided. For the resampling methods, the user can choose between a parametric bootstrap approach (see e.g. Konietschke et al. (2015)) and a Wild bootstrap using Rademacher weights (see e.g. Bathke et al. (2016)).
GUI.RM

A graphical user interface for the RM() function

Description

This function provides a graphical user interface for calculating statistical tests in repeated measures designs.

Usage

GUI.RM()

Details

The function produces a GUI for the calculation of the test statistics and for plotting. Data can be loaded via the "load data" button. The formula, number of resampling iterations (default: 10,000) and the significance level alpha (default: 0.05) need to be specified. Furthermore, the number of sub-plot factors and the column name specifying the subjects in the data need to be provided. For the resampling methods, the user can choose between a permutation approach (Friedrich et al. (2017)), a parametric bootstrap approach (Konietschke et al. (2015)) and a Wild bootstrap using Rademacher weights (Bathke et al. (2016)). If the plot option is chosen, an additional window opens containing information on the plots.
MANOVA function calculates the Wald-type statistic (WTS) and a modified ANOVA-type statistic (MATS) as well as resampling versions of these test statistics for semi-parametric multivariate data.

Usage

MANOVA(formula, data, subject, iter = 10000, alpha = 0.05, resampling = "paramBS", CPU, seed, nested.levels.unique = FALSE, dec = 3)

Arguments

formula A model formula object. The left hand side contains the response variable and the right hand side contains the factor variables of interest. An interaction term must be specified.
data A data.frame, list or environment containing the variables in formula. Data must be in long format and must not contain missing values.
subject The column name of the subjects in the data.
iter The number of iterations used for calculating the resampled statistic. The default option is 10,000.
alpha A number specifying the significance level; the default is 0.05.
resampling The resampling method to be used, one of "paramBS" (parametric bootstrap approach) and "WildBS" (wild bootstrap approach with Rademacher weights).
CPU The number of cores used for parallel computing. If omitted, cores are detected via detectCores.
seed A random seed for the resampling procedure. If omitted, no reproducible seed is set.
nested.levels.unique A logical specifying whether the levels of the nested factor(s) are labeled uniquely or not. Default is FALSE, i.e., the levels of the nested factor are the same for each level of the main factor. For an example and more explanations see the GFD package and the corresponding vignette.
dec Number of decimals the results should be rounded to. Default is 3.

Details

The MANOVA() function provides the Wald-type statistic (WTS) as well as the modified ANOVA-type statistic (MATS) for multivariate designs with metric data as described in Konietschke et al. (2015) and Friedrich and Pauly (2018), respectively. The MATS is invariant under scale transformations of the components and applicable to designs with singular covariance matrices. Both tests
are applicable for non-normal error terms, different sample sizes and/or heteroscedastic variances. They are implemented for designs with an arbitrary number of crossed factors or for nested designs. In addition to the asymptotic p-values, the function also provides p-values based on resampling approaches.

Value

A MANOVA object containing the following components:

- **Descriptive**: Some descriptive statistics of the data for all factor level combinations. Displayed are the number of individuals per factor level combination and the vector of means (one column per dimension).
- **Covariance**: The estimated covariance matrix.
- **WTS**: The value of the WTS along with degrees of freedom of the central chi-square distribution and p-value.
- **MATS**: The value of the MATS.
- **resampling**: p-values for the test statistic based on the chosen resampling approach.

**NOTE**

The number of resampling iterations has been set to 10 in the examples due to run time restrictions on CRAN. Usually it is recommended to use at least 1000 iterations. For more information and detailed examples also refer to the package vignette.

**References**


**See Also**

RM

**Examples**

data(EEG)
EEG_mod <- MANOVA(resp ~ sex * diagnosis, 
               data = EEG, subject = "id", resampling = "paramBS", 
               alpha = 0.05, iter = 10, CPU = 1)
MANOVA.wide

Summary(EEG_mod)

MANOVA.wide  Tests for Multivariate Data in Semi-Parametric Factorial Designs

Description
The MANOVA.wide function calculates the Wald-type statistic (WTS) and a modified ANOVA-type statistic (MATS) as well as resampling versions of these test statistics for semi-parametric multivariate data provided in wide format.

Usage
MANOVA.wide(formula, data, iter = 10000, alpha = 0.05, 
resampling = "paramBS", CPU, seed, nested.levels.unique = FALSE, 
dec = 3)

Arguments

formula  A model formula object. The left hand side contains the matrix of response variables and the right hand side contains the factor variables of interest. An interaction term must be specified.
data  A data.frame, list or environment containing the variables in formula. Data must be in wide format. Note: Lines containing missing values will be removed.
iter  The number of iterations used for calculating the resampled statistic. The default option is 10,000.
alpha  A number specifying the significance level; the default is 0.05.
resampling  The resampling method to be used, one of "paramBS" (parametric bootstrap approach) and "WildBS" (wild bootstrap approach with Rademacher weights). The Wild Bootstrap is calculated for all test statistics.
CPU  The number of cores used for parallel computing. If omitted, cores are detected via detectCores.
seed  A random seed for the resampling procedure. If omitted, no reproducible seed is set.
nested.levels.unique  A logical specifying whether the levels of the nested factor(s) are labeled uniquely or not. Default is FALSE, i.e., the levels of the nested factor are the same for each level of the main factor. For an example and more explanations see the GFD package and the corresponding vignette.
dec  Number of decimals the results should be rounded to. Default is 3.

Value
See MANOVA
NOTE

The number of resampling iterations has been set to 100 in the examples due to run time restrictions on CRAN. Usually it is recommended to use at least 1000 iterations.

See Also

MANOVA

Examples

```r
#Example on producing plastic film from Krzanowski (1998, p. 381), see \code{\link{manova.summary}}
tear <- c(6.5, 6.2, 5.8, 6.5, 6.9, 7.2, 6.9, 6.1, 6.3,
         6.7, 6.6, 7.2, 7.1, 6.8, 7.1, 7.0, 7.2, 7.5, 7.6)
gloss <- c(9.5, 9.9, 9.6, 9.2, 9.1, 10.0, 9.9, 9.4,
         9.1, 9.3, 8.3, 8.4, 8.5, 9.2, 8.8, 9.7, 10.1, 9.2)
opacity <- c(4.4, 6.4, 3.0, 4.1, 0.8, 5.7, 2.0, 3.9, 1.9, 5.7,
           2.8, 4.1, 3.8, 1.6, 3.4, 8.4, 5.2, 6.9, 2.7, 1.9)
rate <- gl(2,10, labels = c("Low", "High"))
additive <- gl(2, 5, length = 20, labels = c("Low", "High"))
example <- data.frame(tear, gloss, opacity, rate, additive)
fit <- MANOVA.wide(cbind(tear, gloss, opacity) ~ rate * additive,
data = example, iter = 100, CPU = 1)
summary(fit)
```

MANOVARM

**MANOVA.RM**: A package for calculating test statistics and their resampling versions for heteroscedastic semi-parametric multivariate data or repeated measures designs.

Description

The MANOVA.RM package provides two important functions: MANOVA() and RM() which will be explained in detail below.

MANOVA and MANOVA.wide function

The MANOVA() and MANOVA.wide() functions provide the Wald-type statistic (WTS) as well as a modified ANOVA-type statistic (MATS) as in Friedrich and Pauly (2018) for multivariate designs with metric data as described in Konietschke et al. (2015). These are applicable for non-normal error terms, different sample sizes and/or heteroscedastic variances. The MATS can even handle designs involving singular covariance matrices. The tests are implemented for designs with an arbitrary number of crossed factors or for nested designs. In addition to the asymptotic p-values, they also provide p-values based on resampling approaches (parametric or wild bootstrap). The difference between the two functions is the format of the data: For MANOVA(), the data needs to be in long format, while MANOVA.wide() is for data in wide format. For further details, see MANOVA and MANOVA.wide.
RM function

The RM() function provides the Wald-type statistic (WTS) as well as the ANOVA-type statistic (ATS) for repeated measures designs with metric data as described in Friedrich et al. (2017). These are even applicable for non-normal error terms and/or heteroscedastic variances. It is implemented for designs with an arbitrary number of whole-plot and sub-plot factors and allows for different sample sizes. In addition to the asymptotic p-values, it also provides p-values based on resampling approaches (Permutation, parametric bootstrap, Wild bootstrap). For further details, see RM.

References


Oxygen Consumption of Leukocytes

Description

A dataset containing measurements on the oxygen consumption of leukocytes in the presence and absence of inactivated staphylococci.

Usage

data(o2cons)

Format

A data frame with 144 rows and 5 variables:

- **O2**  oxygen consumption of leukocytes in µl
- **Staphylococci**  whether or not inactivated staphylococci were added, 1 denotes yes, 0 no
- **Time**  the measurements were taken after 6, 12 and 18 minutes
- **Group**  the treatment group, either P for Placebo or V for Verum
- **Subject**  the subject id
plot.RM

Source


Examples

```r
library(tidyverse)
ggplot(o2cons, aes(x=Group, y=O2)) + geom_point(alpha=0.5) + facet_grid(Staphylococci~Time) +
stat_summary(fun.y = mean, fun.ymin = min, fun.ymax = max, colour = "red")
```

plot.RM

Plot function for an RM object

Description

Generic plot function for RM objects: Returns a plot of the mean values along with confidence intervals for a factor (combination) specified by the user.

Usage

```r
## S3 method for class 'RM'
plot(x, CI.info = FALSE, ...)
```

Arguments

- `x` An object of class RM
- `CI.info` If CI.info = TRUE, the mean values and confidence limits of the considered contrast are printed.
- `...` Additional parameters to be passed to plot()

Details

An additional argument factor can be used to specify the factor(s) used for plotting in two- and higher-way layouts. See the examples for details.
print.MANOVA

Display MANOVA object

Description

Returns a short summary of the results (test statistics with p-values)

Usage

## S3 method for class 'MANOVA'
print(x, ...)

Arguments

x A MANOVA object
...

print.RM

Display an RM object

Description

Returns a short summary of the results (test statistics with p-values)

Usage

## S3 method for class 'RM'
print(x, ...)

Arguments

x An RM object
...

Additional parameters (currently not used)
Description

The RM() function calculates the Wald-type statistic (WTS), the ANOVA-type statistic (ATS) as well as resampling versions of these test statistics for semi-parametric repeated measures designs.

Usage

RM(formula, data, subject, no.subf = 1, iter = 10000, alpha = 0.05, resampling = "Perm", CPU, seed, CI.method = "t-quantile", dec = 3)

Arguments

- **formula**: A model formula object. The left hand side contains the response variable and the right hand side contains the factor variables of interest. An interaction term must be specified. The time variable must be the last factor in the formula.
- **data**: A data.frame, list or environment containing the variables in formula. Data must be in long format and must not contain missing values.
- **subject**: The column name of the subjects in the data. NOTE: Subjects within different groups of whole-plot factors must have individual labels, see Details for more explanation.
- **no.subf**: The number of sub-plot factors in the data, default is 1.
- **iter**: The number of iterations used for calculating the resampled statistic. The default option is 10,000.
- **alpha**: A number specifying the significance level; the default is 0.05.
- **resampling**: The resampling method to be used, one of "Perm" (randomly permute all observations), "paramBS" (parametric bootstrap approach) and "WildBS" (wild bootstrap approach with Rademacher weights). Except for the Wild Bootstrap, all methods are applied to the WTS only.
- **CPU**: The number of cores used for parallel computing. If omitted, cores are detected via detectCores.
- **seed**: A random seed for the resampling procedure. If omitted, no reproducible seed is set.
- **CI.method**: The method for calculating the quantiles used for the confidence intervals, either "t-quantile" (the default) or "resampling" (the quantile of the resampled WTS).
- **dec**: Number of decimals the results should be rounded to. Default is 3.
Details

The RM() function provides the Wald-type statistic as well as the ANOVA-type statistic for repeated measures designs with metric data as described in Friedrich et al. (2017). These are even applicable for non-normal error terms and/or heteroscedastic variances. It is implemented for designs with an arbitrary number of between-subject (whole-plot) and within-subject (sub-plot) factors and allows for different sample sizes. In addition to the asymptotic p-values, it also provides p-values based on resampling approaches. NOTE: The number of within-subject factors needs to be specified in the function call. If only one factor is present, it is assumed that this is a within-subjects factor (e.g. time).

If subjects in different groups of the whole-plot factor have the same id, they will not be identified as different subjects and thus it is erroneously assumed that their measurements belong to one subject. Example: Consider a study with one whole-plot factor "treatment" with levels verum and placebo and one sub-plot factor "time" (4 measurements). If subjects in the placebo group are labelled 1-20 and subjects in the verum group have the same labels, the program erroneously assumes 20 individuals with 8 measurements each instead of 40 individuals with 4 measurements each.

Value

An RM object containing the following components:

Descriptive
Some descriptive statistics of the data for all factor level combinations. Displayed are the number of individuals per factor level combination, the mean and 100*(1-alpha)% confidence intervals (based on t-quantiles).

Covariance
The estimated covariance matrix.

WTS
The value of the WTS along with degrees of freedom of the central chi-square distribution and corresponding p-value.

ATS
The value of the ATS, degrees of freedom of the central F distribution and the corresponding p-value.

resampling
P-values for the test statistics based on the chosen resampling approach.

References


See Also

GFD, nparLD, MANOVA
Examples

data(o2cons)
## Not run:
oxy <- RM(O2 ~ Group * Staphylococci * Time, data = o2cons,
        subject = "Subject", no.subf = 2, iter = 1000, resampling = "Perm", CPU = 1)
summary(oxy)
plot(oxy, factor = "Group")

# For more details including the output of the examples also refer to the # package vignette.

# using the EEG data, consider additional within-subjects factors 'brain region'
# and 'feature'
data(EEG)
EEG_model <- RM(resp ~ sex * diagnosis * feature * region,
        data = EEG, subject = "id", no.subf = 2, resampling = "WildBS",
        iter = 1000, alpha = 0.01, CPU = 4, seed = 987, dec = 2)
summary(EEG_model)

## End(Not run)

---

**simCI**

*Multivariate post-hoc comparisons and simultaneous confidence intervals for contrasts in multivariate factorial designs*

### Description

Multivariate post-hoc comparisons and simultaneous confidence intervals for contrasts in multivariate factorial designs

### Usage

```r
simCI(object, contrast = c("pairwise", "user-defined"), contmat = NULL,
      type = NULL, base = 1, ...)
```

### Arguments

- **object**: A MANOVA object.
- **contrast**: The contrast matrix of interest, can either be "pairwise" or "user-defined".
- **contmat**: If contrast = "user-defined", the contrast matrix must be specified here. Note that its rows must sum to zero.
- **type**: If contrast is "pairwise", the type of the pairwise comparison must be specified here. Calculation is based on the contrMat function in package multcomp, see the corresponding help page for details on the types of contrasts available.
- **base**: an interger specifying which group is considered the baseline group for Dunnett contrasts, see `contrMat`
- **...**: Not used yet.
**Details**

The `simCI()` function computes the multivariate p-values for the chosen contrast of the multivariate mean vector based on the bootstrap version of the sum statistic. Details on this test can be found in Friedrich and Pauly (2018). Furthermore, confidence intervals for summary effects (i.e., averaged over each dimension), also based on the bootstrap version of the sum statistic, are returned as well.

**Value**

Multivariate p-values and simultaneous confidence intervals for the chosen contrasts.

**References**


**See Also**

`contrMat`
Summarizing an RM object

Returns a summary of the results including mean values, variances and sample sizes for all groups as well as test statistics with degrees of freedom and p-values

## S3 method for class 'RM'

```r
summary(object, ...)
```

### Arguments

- **object**: An RM object
- **...**: Additional parameters (currently not used)
Index

*Topic datasets
  EEG, 3
  EEGwide, 4
  o2cons, 11

  conf.reg, 2
  contrMat, 16, 17

  detectCores, 7, 9, 14

  EEG, 3
  EEGwide, 4

  formula, 7, 9, 14

  GFD, 15
  GUI.MANOVA, 5
  GUI.MANOVAwide, 6
  GUI.RM, 6

  MANOVA, 7, 9, 10, 15
  MANOVA.wide, 9
  MANOVARM, 10
  MANOVARM-package (MANOVARM), 10

  nparLD, 15

  o2cons, 11

  plot.RM, 12
  print.MANOVA, 13
  print.RM, 13

  RM, 8, 14

  simCI, 16
  summary.MANOVA, 17
  summary.RM, 18