Package ‘schoRsch’

December 15, 2018

Type Package
Title Tools for Analyzing Factorial Experiments
Version 1.5
Date 2018-12-15
Author Roland Pfister, Markus Janczyk
Maintainer Roland Pfister <roland.pfister@psychologie.uni-wuerzburg.de>
Imports utils
Suggests ez
Description Offers a helping hand to psychologists and other behavioral scientists who routinely deal with experimental data from factorial experiments. It includes several functions to format output from other R functions according to the style guidelines of the APA (American Psychological Association). This formatted output can be copied directly into manuscripts to facilitate data reporting. These features are backed up by a toolkit of several small helper functions, e.g., offering out-of-the-box outlier removal. The package lends its name to Georg "Schorsch" Schuessler, ingenious technician at the Department of Psychology III, University of Wuerzburg. For details on the implemented methods, see Roland Pfister and Markus Janczyk (2016) <doi: 10.20982/tqmp.12.2.p147>.

URL http://www.tqmp.org/RegularArticles/vol12-2/p147/index.html
License GPL
NeedsCompilation no
Repository CRAN
Date/Publication 2018-12-15 22:20:04 UTC

R topics documented:

anova_out ................................................................. 2
cd ................................................................. 3
chi_out ................................................................. 4
clear ................................................................. 5
cor_out ................................................................. 5
ntiles ................................................................. 6
Format ANOVA Output

Description
Distills the most relevant data from an output object of ezANOVA and displays the results in a compact format.

Usage
anova_out(ezout, print = TRUE, sph.cor = "GG", mau.p = 0.05, etasq = "partial", dfsep = " ", corr.df = FALSE, show.eps = 0)

Arguments
ezout  Output object created by a call to ezANOVA. This call has to have included a detailed output (default=TRUE).
print  Force results to be displayed, even if the function output is assigned to a variable (e.g., output <- anova_out(...)); logical; default=TRUE).
sph.cor Correction method (one of "no","GG","HF"; default="GG").
mau.p  Threshold for Mauchly’s test of sphericity (numerical; default=0.05).
etasq  Effect size estimate to be used; either partial eta-squared("partial"; default) or generalized eta-squared("generalized").
dfsep  String that delimits the degrees of freedom of each F-value (default=" ", ").
corr.df Display corrected degrees of freedom when Mauchly’s test of sphericity is significant (default=FALSE).
show.eps Show epsilon estimates when Mauchly’s test of sphericity is significant? 0 = do not show, 1 = print after denominator dfs, 2 = print after F-value, 3 = print after effect size (default=0).

Details
The output of a call to ezANOVA is formatted according to the guidelines of the APA (American Psychological Association) as well as the DGP's ("Deutsche Gesellschaft fuer Psychologie"; German Psychological Society).

For repeated-measures ANOVAs, sphericity corrections are automatically applied to the p-values of effects that produced a significant result in Mauchly’s test of sphericity. The corresponding input arguments (sph.cor and mau.p) do not affect between-subject designs.
cd

Value
anoout(ezout,...) returns a list containing (1) the ANOVA table, (2) sphericity tests and corrections (if applicable), (3) formatted ANOVA results, (4) notes about which correction was applied to which effect.

Author(s)
Roland Pfister, Markus Janczyk

See Also
ezANOVA; aov; chi_out; cor_out; t_out;

description
Change Directory

Description
Performs relative changes of the working directory. Calling cd("...") moves one level up in the hierarchy whereas cd("folder_name") moves one level down to the designated folder.

Usage
cd(x)

Arguments
x : A character string corresponding to target directory or "..".

Details
cd is designed as an equivalent to the DOS command. Contrary to the common use of cd, however, this function does not take absolute paths as input. Use setwd instead to navigate to an absolute path.

Value
cd(x) returns the new working directory.

Author(s)
Roland Pfister, Markus Janczyk

See Also
getwd; setwd;
### Examples

```r
## Create temporary folder
dir.create("a_test_dir")

## Navigate into the new folder...
\texttt{cd("a_test_dir")}
## ... and back again
\texttt{cd("..")}

## Remove temporary folder
unlink("a_test_dir", recursive=TRUE)
```

---

### Description

Distilles the most relevant data from an output object of chisq.test and displays the results in a compact format.

### Usage

```r
\texttt{chi_out(chioutput, show.n = FALSE, print = TRUE)}
```

### Arguments

- **chioutput**: Output object created by a call to chisq.test.
- **show.n**: Display sample size (logical; default=FALSE)
- **print**: Force results to be displayed, even if the function output is assigned to a variable (e.g., output \texttt{<- chi_out(...)}; logical; default=TRUE).

### Details

The output of a call to \texttt{chi_out} is formatted according to the guidelines of the APA (American Psychological Association) as well as the DGPs ("Deutsche Gesellschaft fuer Psychologie"; German Psychological Society).

### Value

\texttt{\textit{chi_out(chioutput,\ldots)}} returns a data.frame containing (1) a description of the test and (2) a line with formatted results.

### Author(s)

Daniel Gromer
clear

See Also

chisq.test, anova_out, cor_out, t_out;

clear Clear Global Workspace

Description

The global workspace is cleared; clear is a shortcut for the usual `rm(list=ls())`.

Usage

clear()

Author(s)

Roland Pfister, Markus Janczyk

See Also

rm, ls;

cor_out Format Correlation Test Statistics

Description

Distilles the most relevant data from an output object of cor.test and displays the results in a compact format.

Usage

cor_out(coroutput, stats = FALSE, print = TRUE)
Arguments

- **coroutput**: Output object created by a call to `cor.test`.
- **stats**: If TRUE, the output includes t-values and corresponding degrees of freedom (default=FALSE).
- **print**: Force results to be displayed, even if the function output is assigned to a variable (e.g., `output <- cor_out(...)`; logical; default=TRUE).

Details

The output of a call to `cor.test` is formatted according to the guidelines of the APA (American Psychological Association) as well as the DGPs ("Deutsche Gesellschaft fuer Psychologie"; German Psychological Society).

Value

cor_out(coroutput,...) returns a line containing the formatted correlation results.

Author(s)

Markus Janczyk, Roland Pfister

See Also

cor; cor.test; anova_out; chi_out; t_out;

Description

The data of a variable are rank-ordered and split to bins of (approximately) equal size. When tied ranks span across category borders, the function assigns all values to the lowest possible bin. This procedure can result in slightly different results as the corresponding function Rank Cases of SPSS with option Ntiles.

Usage

```r
ntiles(data, dv,
    factors = NaN,
    bins = 5,
    res.labels = FALSE)
```
ntiles  7

Arguments

data   A data frame containing the data relevant variable and possible factors that can be used to split the data frame into separate compartments.
dv     Character string specifying the name of the variable within data that is to be cut in bins. Alternatively, dv can be the appropriate column index.
factors A string or vector of strings (e.g., c("subject","condition")) stating the conditions that should be used for splitting the data.
bins   The number of bins to be generated. Alternatively, a vector of cut-points can be specified according to the break argument of the function cut.
res.labels The default value FALSE returns the bin number for each observation whereas TRUE returns the corresponding interval borders (in ranks).

Value

ntiles(data, dv, ...) returns a vector of bins.

Author(s)

Roland Pfister; Markus Janczyk

See Also

cut; rank; split; lapply;

Examples

## Build data frame
var1 <- c(1:9)
var2 <- c(1,1,1,2,2,2,3,3,3)
tmpdata <- data.frame(cbind(var1, var2))
tmpdata$var2 <- as.factor(tmpdata$var2)

## Get overall bins and display result
tmpdata$bins <- ntiles(tmpdata, dv = "var1", bins=3)
tmpdata

## Get bins separately for each factor level
## and display result
##
tmpdata$bins2 <- ntiles(tmpdata, dv = "var1", bins=3, factors = "var2")
tmpdata
Screen Data for Outliers

Description

A chosen column of a data frame is screened for outliers, outliers are marked and/or eliminated. Either absolute lower and upper limits are applied, or outliers are identified based on z-transformed data. Either exact limits and/or cutoffs based on z-values need to be entered.

Usage

```r
outlier(data, dv,  
todo = "na", res.name = "outlier",  
upper.limit = NaN, lower.limit = NaN,  
limit.exact = FALSE,  
upper.z = NaN, lower.z = NaN,  
z.exact = FALSE, factors = NaN,  
z.keep = TRUE, z.name = "zscores",  
vsj = FALSE,  
print.summary = TRUE)
```

Arguments

data A data frame containing the data to be screened as well as appropriate condition variables.
dv Character string specifying the name of the variable within data that is to be screened for outlier. Alternatively, dv can be the appropriate column index.
todo Character string specifying the fate of outliers: "na" - outliers are turned into NAs, "elim" - rows containing outliers are deleted from dataframe, "nothing" - nothing happens, default=todo = "na".
res.name Character string specifying the name of the variable to be used for marking outliers, default=res.name = "outlier".
upper.limit An optional numerical specifying the absolute upper limit defining outliers.
lower.limit An optional numerical specifying the absolute lower limit defining outliers.
limit.exact Logical, if TRUE values equal to lower.limit/upper.limit are deemed outlier.
upper.z An optional numerical specifying how much standard deviations within a cell a value must exceed to be identified as an outlier.
lower.z An optional numerical specifying how much standard deviations within a cell a value must undercut to be identified as an outlier.
factors A string or vector of strings (e.g., c("subject","condition")) stating the conditions that should be used for splitting the data.
z.exact Logical, if TRUE z-values equal to lower.z/upper.z are deemed outlier.
z.keep Logical, if TRUE, z-scores are stored in an additional column. If FALSE, z-scores are discarded after outlier correction.
**Details**

If both, absolute limits and z-limits are specified, absolute limits are processed first and z-scores are computed for the remaining data points.

**Value**

`outlier(data,...)` returns the original data frame with the outlier correction applied. This data frame also has one additional column containing flags for outliers (0 = not suspicious, 1 = outlier). If z-scores are requested, these scores are returned as an additional column.

**Author(s)**

Markus Janczyk, Roland Pfister

**See Also**

`split`; `zscores`;

---

**Description**

Offers a helping hand to psychologists and other behavioral scientists who routinely deal with experimental data from factorial experiments. It includes several functions to format output from other R functions according to the style guidelines of the APA (American Psychological Association). This formatted output can be copied directly into manuscripts to facilitate data reporting. These features are backed up by a toolkit of several small helper functions, e.g., offering out-of-the-box outlier removal. The package lends its name to Georg "Schorsch" Schuessler, ingenious technician at the Department of Psychology III, University of Wuerzburg.

**Details**

<table>
<thead>
<tr>
<th>Package</th>
<th>schoRsch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Package</td>
</tr>
<tr>
<td>Version</td>
<td>1.5</td>
</tr>
<tr>
<td>Date</td>
<td>2018-12-15</td>
</tr>
<tr>
<td>License</td>
<td>GPL-3</td>
</tr>
</tbody>
</table>
This package contains the following functions:

- **anova_out**: Formats the output object from `ezANOVA` to the APA style (requires the `ez` package).
- **cor_out**: Formats the output object from `cor.test` to the APA style.
- **chi_out**: Formats the output object from `chisq.test` to the APA style.
- **t_out**: Formats the output object from `t.test` to the APA style.
- **outlier**: Screens data for outliers, based on absolute values or z-scores. Outliers can either be marked or eliminated.
- **ntiles**: Split distribution into quantiles for distribution analysis.
- **zscores**: Computes z-scores of values separately for defined design cells.
- **cd**: To easily change the current working directory.
- **toclipboard**: Write data to clipboard (Windows only).
- **clear**: Clears the whole workspace (i.e., like `rm(list=ls())`)

Version history:

- **v1.5** | 2018-12-15 | Default value for correcting effect sizes for paired-samples t-tests changed to `FALSE`; the use of corrections is now displayed as feedback message. Also: New options for `anova_out`; it is now possible to display corrected degrees of freedom for violations of the sphericity assumption and corresponding epsilon estimates. Thanks to Onur Asci for the feature request. Additional bugfix for the dfsep argument of `anova_out`.
- **v1.4** | 2017-02-14 | Bugfix for `cor_out` that no longer displays leading zeros for correlation coefficients; thanks to Juan Ramon Barrada for sending in the bug report.
- **v1.3** | 2016-09-13 | Overall documentation update based on comments from Vincent LeBlanc.
- **v1.2** | 2015-07-05 | Bugfix for the `print` option of `anova_out`; thanks to Sylvain Clement for sending in the bug report. Minor code changes.
- **v1.1** | 2014-07-30 | New functions `chi_out` (contributed by Daniel Gromer) and `toclipboard`; bugfixes when `anova_out` is called without `detailed=TRUE`. Updated help files.
- **v1.0** | 2013-03-20 | Package release.

**Author(s)**

Roland Pfister <roland.pfister(at)psychologie.uni-wuerzburg.de>, Markus Janczyk;

**References**

Description

A data frame of variable is written to the clipboard, allowing easy pasting to MS Excel and Open/Libre Office Calc. This function is a wrapper to `write.table` with pre-specified options for plug-and-play usage. Most options of `write.table` are also supported by `toclipboard`. Note: The current version of `toclipboard` only supports Windows systems; the function will not run under Linux or Mac OS.

Usage

```r
toclipboardHdataL
    sep = "\t", quote = FALSE,
    eol = "\n", na = "NA",
    dec = ".", row.names = FALSE,
    col.names = TRUE)
```

Arguments

- **data**: The first argument should be the data frame or variable that is to be written to the clipboard. Data frames are copied with column names but without row names and columns are separated by tabs. This behavior can be customized with the following optional arguments (passed to `write.table`).
- **sep**: Delimiter string.
- **quote**: Put quotes around strings?
- **eol**: End-of-line character.
- **na**: How should NA-values be written?
- **dec**: Decimal separator.
- **row.names**: Should row names be written?
- **col.names**: Should column names be written?

Author(s)

Roland Pfister

See Also

- `write.table`
Examples

```r
## Build data frame
t_out <- c(1:9)
t2_out <- c(1,1,1,1,2,2,2,3,3)
tmpdata <- data.frame(cbind(t_out,t2_out))

## Write data frame to clipboard
toclipboard(tmpdata)
```

### Format t-Test Output

**Description**

Distill the most relevant data from an output object of `t.test` and displays the results in a compact format.

**Usage**

`t_out(toutput, n.equal = TRUE, welch.df.exact = TRUE, welch.n = NA, d.corr = FALSE, print = TRUE)`

**Arguments**

- `toutput` Output object created by a call to `t.test`.
- `n.equal` Only applicable to two-sample t-tests. If sample sizes are not equal, `n.equal` specifies a vector of sample sizes, e.g., `n.equal = c(12,8)`.
- `welch.df.exact` Only applicable to Welch-tests. Indicates whether Welch-adjusted or unadjusted degrees of freedom (dfs) are reported (default=TRUE, i.e., Welch-adjusted dfs). If set to FALSE, the parameter `welch.n` has to be set as well.
- `welch.n` Only applicable to Welch-tests with unadjusted degrees of freedom. Parameter should be equal to the total sample size `n=n_1+n_2`.
- `d.corr` Only applicable to one-sample or paired-samples t-tests. If TRUE, Cohen’s ds are computed using `sqrt(2)`-correction. Default changed to FALSE from version 1.5 onwards with an additional feedback message showing the use of corrections.
- `print` Force results to be displayed, even if the function output is assigned to a variable (e.g., `output <- t_out(...)`; logical; default=TRUE).

**Details**

The output of a call to `t_out` is formatted according to the guidelines of the APA (American Psychological Association) as well as the DGPs ("Deutsche Gesellschaft fuer Psychologie"; German Psychological Society).
zscores

**Value**

t_out(output,...) returns a list containing (1) a description of the t-test (two-sample t-test, Welch-test, paired-samples t-test, one-sample t-test) and (2) a line with formatted results.

**Author(s)**

Roland Pfister, Markus Janczyk

**See Also**

t.test; anova_out; chi_out; cor_out;

---

**Description**

Data of an input vector is transformed to z-scores (mean = 0, sd = 1). The function operates on single vectors as well as on specified columns of a data frame.

**Usage**

zscores(data, factors=NaN, dv=NaN)

**Arguments**

data Either a data frame containing the data of interest or a single vector.

factors If called with factors=NaN (default), the entire data is processed according to its grand mean and total variance. If data is a vector, factors can be a list of variables for splitting the variable into separate compartments. If data is a data frame, factors has to be specified as a character vector of column names or column indices.

dv If data is a single vector, dv does not have to be specified. If data is a data frame, dv indicates the column of the data frame which contains the variable for z-transformation (e.g., dv="rt") or its column index (e.g., dv=15).

**Details**

zscores computes z-score of a vector or a specified column within a data frame. Computation can be done separately for combinations of factors.

**Value**

zscores() returns a vector containing the requested z-scores.

**Author(s)**

Roland Pfister, Markus Janczyk
See Also

scale; split; outlier:

Examples

# Create input vector and compute z-scores
measurements <- c(3,12,5,4,2,23,1,6)
zscores(measurements)

# Compute z-scores separately
# for conditions
cond1 <- c(1,1,1,2,2,2,2)
cond2 <- c(1,1,2,2,1,1,2,2)
zscores(measurements,list(cond1))
zscores(measurements,list(cond1,cond2))

# Calling zscores for data frames
data <- data.frame(measurements,
cond1,cond2)
zscores(data,dv="measurements",
factors=c("cond1","cond2"))

# Operating on column indices
zscores(data,dv=1,
factors=3)
Index

+Topic **package**
  schorsch, 9
+Topic **utilities**
  cd, 3
  clear, 5
  toclipboard, 11
  zscores, 13
  anova_out, 2, 5, 6, 10, 13
  aov, 3
  cd, 3, 10
  chi_out, 3, 4, 6, 10, 13
  chisq.test, 5, 10
  clear, 5, 10
  cor, 6
  cor.test, 6, 10
  cor_out, 3, 5, 5, 10, 13
  cut, 7
  ez, 10
  ezANOVA, 3, 10
  getwd, 3
  lapply, 7
  ls, 5
  ntiles, 6, 10
  outlier, 8, 10, 14
  rank, 7
  rm, 5
  scale, 14
  schoRsch, 9
  schoRsch-package (schoRsch), 9
  setwd, 3
  split, 7, 9, 14
  t.test, 10, 13
  t_out, 3, 5, 6, 10, 12
  toclipboard, 10, 11
  write.table, 11
  zscores, 9, 10, 13