Package 'bmisc'

August 15, 2011

Type Package

Version 0.2-12

Title Miscellaneous functions

| Author Benoit Bruneau | |
|---|---|
| Maintainer Benoit Bruneau Senoit.bruneau1@gmail.com> | |
| Description This package has different functions that I have accumulated with time. This is the Alpha version. | |
| Depends car, lattice, zoo, robustbase, methods | |
| License LGPL $>= 3.0$ | |
| R topics documented: | |
| get.partial.etas | 2 4 5 5 6 7 8 9 9 10 11 |

19

2 att.strp

| n | | | 21 | 1 |
|------------|-------------------------------------|------|--------|---|
| | $	ext{m.test}$ | | | |
| P.a | djust | | 25 | |
| | k.list | | | |
| | :.diff | | | 6 |
| - | formance | | | 7 |
| | .logit | | | 8 |
| * . | 5.ypr | | | 8 |
| | | | 29 | 9 |
| _ | $\operatorname{lors} \ldots \ldots$ | | | 0 |
| | ct.z | | | 0 |
| rep | ace.z | | 31 | 1 |
| res | d.ortho | | 31 | 1 |
| riv | rd | | 31 | 1 |
| $_{ m rm}$ | levels | | 32 | 2 |
| rol | min | | 32 | 2 |
| rou | ndup | | | 3 |
| rur | max | | | 3 |
| rur | mean | | 34 | 4 |
| rur | min | | 34 | 4 |
| s.a | L | | 34 | 4 |
| se | | | 35 | 5 |
| sho | w.North | | 35 | 5 |
| sor | .vdf | | | 6 |
| sur | nmary.ypr | | 37 | 7 |
| tte | t.perm | | 37 | 7 |
| un | oad | | 39 | 9 |
| we | k.1 | | 39 | 9 |
| we | k.num | | 40 | 0 |
| yp | .1 | | 40 | 0 |
| Index | | | 44 | 4 |

 ${\it att.strp} \qquad \qquad {\it Attibute \ stripper}$

Description

Strips an object of its attributes

Usage

att.strp(x)

Arguments

x the name of an object (vector, matrix, data.frame, array or list)

att.strp 3

Details

This function strips an object of its attributes. In the case of a vector, all attributes are removed. For a matrix or an array, only c('dim', 'dimnames') are kept. When att.strp is used on a data.frame, all attributes of the variables are striped and only c('names', 'row.names', 'na.action', 'class') are kept for the data.frame object.

Value

returns an object of the same class as the original one.

Author(s)

Benoit Bruneau

```
Creating different objects
    with added attributes (label)
### numerical vector ###
x <- 1:10
attr(x,"label") <- "test1"
attributes(x)
### data frame ###
z=data.frame(x,x)
attr(z,"labels") <- "test2"</pre>
attributes(z)
attributes(z[,1])
attributes(z[,2])
### array ###
y=array(x,c(2,2,2))
attr(y,"labels") <- "test3"</pre>
attributes(y)
attributes(y[,,1])
attributes(y[,,2])
### list containing the vector, ###
### data frame and array
u=list(x,z,y)
attr(u,"labels") <- "test4"
attributes(u)
attributes(u[[1]])
attributes(u[[2]])
attributes(u[[3]])
attribute stripping
x2=att.strp(x)
z2=att.strp(z)
y2=att.strp(y)
```

4 bmisc

```
u2=att.strp(u)
verification of the attributes
    for all stripped objects
### numerical vector ###
attributes(x2)
### data frame ###
attributes(z2)
attributes(z2[,1])
attributes(z2[,2])
### array ###
attributes(y2)
attributes(y2[,,1])
attributes(y2[,,2])
### list containing the vector, ###
### data frame and array
                             ###
attributes(u2)
attributes(u2[[1]])
                       # vector in the list
attributes(u2[[2]])
                       # data frame in the list
attributes(u2[[2]][,1])
                       # data frame in the list
                       # data frame in the list
attributes(u2[[2]][,2])
attributes(u2[[3]]
                       # array in the list
attributes(u2[[3]][,,1]) # array in the list
attributes(u2[[3]][,,2]) # array in the list
```

bmisc

 $Miscellaneous\ functions$

Description

This package has different functions that I have accumulated with time. I am not the author of all of them even though I have modified most of them. This is the Alpha version.

Format

 $\begin{array}{lll} \text{Package:} & \text{bmisc} \\ \text{Type:} & \text{Package} \\ \text{Version:} & 0.2\text{-}12 \\ \text{Date:} & 04\text{-}08\text{-}2011 \\ \text{License:} & \text{LGPL} >= 3.0 \\ \end{array}$

ceiling.lg 5

Details

For pdf version of the help, write vignette("bmisc").

Author(s)

Benoit Bruneau

Maintainer: Benoit Bruneau

 denoit.bruneau1@gmail.com>

ceiling.lg

 $ceiling\ largest$

Description

Ceiling to largest digit

Usage

```
ceiling.lg(x)
```

Arguments

х

Numeric vector

Details

Gives the ceiling to largest digit (i.e., $54 \rightarrow 60$).

Examples

```
ceiling.lg(250)
ceiling.lg(25000000)
```

clean

 $Clean\ a\ Data\ Frame$

Description

Cleans a data.frame from a starting point with a defined threshold

Usage

```
clean(data= x, col.start =1, min.val=NULL)
```

Arguments

data then name of the data.frame

col.start indicate the columns from which to start reading

min.val numeric. Read details

6 corr.perm

Details

min.val is the minimum value accepted in a column. Colomns with this value or higher will be kept in the data.frame.

More will be added to this function.

Value

returns the data.frame with the clean columns

Author(s)

Benoit Bruneau

Examples

```
x=rnorm(50 , 20, 12)
y=runif(50 )
z=rpois(50, 3)
v=x*y/z
t=z*v
pp=data.frame(aa=x, bb=y, cc=v, dd=z, ee=t)
summary(pp)
pp1 = clean(pp, min.val=0.06)
```

corr.perm

Pearson Correlation by Permutation

Description

Tests the Pearson correlation estimate (r) by use of permutation

Usage

```
corr.perm(x,y,nperm=999)
```

Arguments

x,y Two vectors of same length used for correlation analysis

nperm Number of permutations (default = 999)

Value

P.perm pvalue estimated by permutations

P.para parametric pvalue estimated

inf inferior limit of the confidence interval sup superior limit of the confidence interval

df degree of freedom

cv 7

Examples

```
x <- rnorm(50,0,1)
y <- runif(50,0,1)*x
toto = corr.perm(x, y)</pre>
```

cv

Coefficient of Variation (CV)

Usage

```
cv(x, na.rm=T)
```

Arguments

x an R object (vector, matrix,...)

na.rm a logical value indicating whether NA values should be stripped before

the computation proceeds

Details

The coefficient of variation (CV) is the ratio of the standard deviation to the mean. The CV is defined for the absolute value of the mean to ensure it is always positive.

Examples

```
x=rnorm(50)
cv(x)
```

day

day

Description

Day of year as decimal number (001-366).

Usage

day(x)

Arguments

x

```
# will soon be available
```

8 Errbar

| Errbar erre | or bars |
|-------------|---------|
|-------------|---------|

Description

Adds error bars on a plot

Usage

Arguments

| х | numeric vector |
|------------|--|
| У | numeric vector |
| xinf, xsup | numeric vectors containing the upper (xsup) and/or lower (xinf) limits of the confidence interval for x-axis values. |
| yinf, ysup | numeric vectors containing the upper (ysup) and/or lower (yinf) limit of the confidence interval for y-axis values. |
| xCI | numeric vectors containing the confidence intervals for x-axis values. |
| yCI | numeric vectors containing the confidence intervals for y-axis values. |
| • • • | additional graphical arguments (par) such as col, lty, lwd and/or arguments for $arrows$. |

Details

If xCI and/or yCI are defined, individually defined limits (ie. xinf, xsup, yinf, ysup) are not used.

See Also

```
arrows, par
```

```
x <- 1:10
y <- x + rnorm(10)

yci <- runif(10)
xci <- runif(10)

plot(x,y, ylim=c(min(y-yci),max(y+yci)))
Errbar( x, y, yCI=yci)

plot(x,y, xlim=c(min(x-xci),max(x+xci)))
Errbar( x, y, xCI=xci )

plot(x,y, ylim=c(min(y-yci),max(y+yci)), xlim=c(min(x-xci),max(x+xci)))
Errbar( x, y, yCI=yci, xCI=xci )

# Gives an Error message</pre>
```

fct 9

```
#plot(x,y, ylim=c(min(y-yci),max(y+yci))) ## adds the yCI and gives
#Errbar( x, y, ysup=1, yCI=yci) ## an error message for the ysup
```

fct

 $Print\ bmisc\ functions$

Description

Print all functions of bmisc package

Usage

fct()

format.hms

Format seconds into hours

Description

Transforms time format

Usage

format.hms(sec)

Arguments

sec

time expressed in seconds $\,$

Value

returns hrs:min:sec

Examples

format.hms(20000)

10 gam.Check

gam.Check

Some diagnostics for a fitted gam model

Description

Takes a fitted gam object produced by gam() and produces some diagnostic information about the fitting procedure and results. The default is to produce 4 residual plots, and some information about the convergence of the smoothness selection optimization.

Usage

Arguments

```
b a fitted gam object as produced by gam().

main a character vector containing the four titles to be used.

xlab a character vector containing the four x labels to be used.

ylab a character vector containing the four y labels to be used.

text a character or expression vector specifying the text to be written.

args.histplot list of additional arguments to pass to histplot()

... additional text and graphical parameters (see par, mtext)
```

Details

This function plots 4 standard diagnostic plots, and some other convergence diagnostics. Usually the 4 plots are various residual plots. The printed information relates to the optimization used to select smoothing parameters. For the default optimization methods the information is summarized in a readable way, but for other optimization methods, whatever is returned by way of convergence diagnostics is simply printed.

This is a modified version of gam.check from mgcv-package so that main titles, x labels and y labels can be customized.

References

Wood S.N. (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC Press.

get.partial.etas 11

Examples

```
library(mgcv)
set.seed(0)
dat <- gamSim(1,n=200)
b<-gam(y~s(x0)+s(x1)+s(x2)+s(x3),data=dat)
plot(b,pages=1)

gam.check(b)
gam.check(b, main=c("A","B","C","D"))</pre>
```

get.partial.etas

get partial etas

Usage

```
get.partial.etas(model)
```

Arguments

model

Examples

will soon be available

histplot

histplot

Usage

Arguments

dat one of:

- a numeric vector
- an object of class c('norm', 'lm', 'aov', 'glm', 'gam') resulting from a calls to c(norm.test,lm,aov,glm,gam)

breaks one of:

12 inv.pred

• a vector giving the breakpoints between histogram cells,

- a single number giving the number of cells for the histogram,
- a character string naming an algorithm to compute the number of cells (see 'Details'),
- a function to compute the number of cells.

In the last three cases the number is a suggestion only.

barc a color to be used to fill the bars.

borc a color to be used for the borders the bars.

fit.norm a logical variable indicating whether to fit a normal density curve (TRUE)

or not (FALSE).

lcol color of the normal density curve

stat the statistic to add on the graph. One of (c("all", "mean", "median")).

Default is NULL.

stat.lab a character vector with the labels for the estimated mean and/or median.

Default is c("Mean", "Median").

rug a logical variable indicating whether to superpose a rug (TRUE) or not

(FALSE).

main the main title of the graph

... additional arguments to be passed to plot (see par)

Details

The default for breaks is "Sturges": see nclass.Sturges. Other names for which algorithms are supplied are "Scott" and "FD" / "Freedman-Diaconis" (with corresponding functions nclass.scott and nclass.FD). Alternatively, a function can be supplied which will compute the intended number of breaks as a function of x.

See Also

hist

Examples

```
x=rnorm(50)
histplot(x)
norm.x=norm.test(x)
histplot(norm.x)
```

inv.pred

 $Inverse\ Predictions\ with\ SE$

Usage

```
inv.pred( object, cf=1:2, y )
```

is.even 13

Arguments

object an object of class c('lm', 'glm') resulting from a calls to c(lm, glm)

cf the linear coefficients ('intercept', 'slope') to be used.

y the y value for which x will be estimated with it's standard error.

Details

More to come.

Author(s)

Benoit Bruneau

is.even

is even

Description

Identifies if a value is even or not

Usage

is.even(x)

Arguments

X

numeric vector

Details

Will returns TRUE if roundup(x) is an even number.

Value

logical

See Also

is.odd

Examples

```
is.even(5)
```

is.even(6)

14 last

is.odd

 $is \ odd$

Description

Identifies if a value is odd or not

Usage

is.odd(x)

Arguments

X

numeric vector

Details

Will returns TRUE if roundup(x) is an odd number.

Value

logical

See Also

is.even

Examples

is.odd(5)
is.odd(6)

last

last

Usage

last(x)

Arguments

х

Examples

will soon be available

lev 15

|--|

Description

Tests heteroscedasticity after an Anova

Usage

```
lev(y, ...)
## S3 method for class 'formula'
lev(y, data=NULL, ...)
## S3 method for class 'lm'
lev(y, ...)
## Default S3 method:
lev(y, group, data=NULL , trim.alpha = 0.1, type="abs",...)
```

Arguments

| у | response variable for the default method, lm class object for the lm method or formula class object for the formula methode. If y is a linear-model object or a formula, the variables on the right-hand-side of the model must all be factors and must be completely crossed. See details. |
|------------|---|
| group | for the default method, factor (concatenated factor when multiple factors). See details. |
| data | data.frame where the dependant variable and the factor(s) are |
| trim.alpha | Alpha level (percentiles) trimming the data on which the mean will be evaluated |
| type | Type of transformation made on the residuals. Either "abs" for absolute values or "sq" for sqared values |
| | |

Details

When using the lm method, data doesn't need to be defined. When using the formula or default methods, data can be defined if the data used is in a data.frame.

When group is manually defined in the default method, use paste(x,y,z) or $\$ interaction(x,y,z) form where "x", "y" and "z" are the factors. There is no restrictions on the number of factors.

O'Brien's (1981) performs test for equality of variances within each group: based on transforming each observation in relation to its group variance and its deviation from its group mean; and performing an ANOVA on these transformed scores (for which the group mean is equal to the variance of the original observations). The procedure is recognised to be robust against violations of normality (unlike F-max).

16 lib.code

Value

Model The model

Levene Results for Levene's test

 ${\tt LeveneTrimMean}$

Results for Levene's test on the trimmed mean

Brown.Forsythe

Results for Brown-Forsythe's test

OBrien Results for O'Brien's test

See Also

```
leveneTest from {car}
```

Examples

lib.code

Retreives the code for lib().

Description

Will print in the R windows the code for lib() (READ DETAILS).

Usage

```
lib.code()
lib(pack, install=TRUE, load=TRUE, quietly=TRUE,
    warn.conflicts=FALSE)
```

Arguments

pack

Character vector specifying which package(s) to load/install.

lsmean 17

Details

USE lib.code() TO GET THE CODE FOR THE FUNCTION lib().

lib.code() prints in R the code for lib(). Copy and paste the code for lib() in the file "C:/Program Files/R/R-2.12.1/etc/Rprofile.site" (Windows) or "~/.Rprofile" (Mac).

lib() will load packages named in a charcater vector. If install is TRUE, packages not yet installed will be installed.

Author(s)

Benoit Bruneau

Examples

lib.code()

lsmean

Least Squares Means

Description

THIS FUNCTION IS FROM PACKAGE pda THAT IS STILL UNDER CONSTRUCTION ON R-Forge. IT HAS BEEN INCLUDED IN bmisc FOR PRACTICAL REASONS.

Caution: This routine is not fully tested for models with nested factors or mixed models. Please check results against another package (e.g. SAS proc mixed). It appears to correctly handle lme objects, but does not work well for aov objects that include $ext{Error}$ type nesting in the formula. Further, it does not properly handle polynomial terms—only the linear term is included. For now, create dummies like $ext{x2} = ext{x*}$ manually and include $ext{x2}$ in your model.

Usage

```
lsmean(object, ...)
## Default S3 method:
lsmean(object, ..., factors, effects = FALSE, se.fit = TRUE,
    adjust.covar = TRUE)
## S3 method for class 'lm'
lsmean(object, data, factors, expr, contrast, effects = FALSE,
    se.fit = TRUE, adjust.covar = TRUE, pdiff = FALSE,
    reorder = FALSE, lsd, level = .05, rdf, coef, cov, ...)
## S3 method for class 'lme'
lsmean(object, data, factors, ..., rdf, coef, cov)
## S3 method for class 'lmer'
lsmean(object, data, factors, expr, ..., rdf, coef, cov)
## S3 method for class 'listof'
lsmean(object, data, factors, stratum, expr, contrast, ...)
```

18 lsmean

Arguments

object response vector (default) or model object (lm).
... factors and covariates (must be same length as y).

data frame in which to interpret variables (found from object if missing).

factors character vector containing names of x.factor and trace.factoras first two

entries. Must be in names(data) and labels(object). Default is all fac-

tor names.

effects drop intercept if TRUE (only works properly with sum-to-zero contrasts).

se.fit compute pointwise standard errors if T.

adjust.covar adjust means to average covariate values if T; otherwise use covariate

mean for each combination of factors.

pdiff Include letters to signify significant differences.

reorder Reorder means from largest to smallest.

lsd Include average LSD if TRUE (also need pdiff=TRUE).

level Significance level for pdiff calculations.

rdf Residual degrees of freedom.

coef Coefficients for fixed effects in object.
cov Covariance matrix for fixed effects.

expr Call expression (formula)

contrast Type of contrasts (default is attribute contrasts of object) stratum Name of stratum for Ismean calculation as character string.

Value

Data frame containing unique factor levels of factors, predicted response (pred) and standard errors (se). WARNING: Ismean may not function properly if there are empty cells. Standard errors for mixed models using methods lmer and listof are not fully debugged.

Author(s)

Brian S. Yandell

See Also

```
predict.
```

```
## Not run:
lsmean(y,x1,x2)
# the following does the same thing
fit <- lm(y~x1+x2)
data <- data.frame(y,x1,x2)
lsmean(fit,data,factors=c("x1","x2")
## End(Not run)</pre>
```

make.z

```
make.z make.z
```

Usage

```
make.z(x, index = NULL)
```

Arguments

X

index

Examples

will soon be available

mc.long

Pairwise t tests in long format

Description

Calculate pairwise T tests between group levels with corrections for multiple testing presented in long format

Usage

Arguments

| У | response variable for the default method, or lm or formula object. If y is a linear-model object or a formula, the variables on the right-hand-side of the model must all be factors and must be completely crossed. |
|----------------|--|
| group | for the default method, factor (concatenated factor when multiple factors). See details. |
| data | data.frame where the dependant variable and the factor(s) are |
| p.adjust.metho | od |
| | method for adjusting p values. Default is Holm's method. (see P.adjust) |
| column | new names for the factor(s); this is optional |
| digits | controls the number of digits for the presented results presented |
| silent | a logical variable indicating whether to indicate the general warning (FALSE) or not (TRUE). |
| | $additional\ arguments\ to\ pass\ to\ {\tt P.adjust}, {\tt pairwise.t.test}\ and/or\ {\tt t.test}.$ |

20 mse

Details

When making multiple t tests for all combinations, the n option of P.adjust can be used to identify the number of comparisons that are actually used. This is only to simplify the uses p values corrections on the full output matrix when only some of the comparisons are meaningfull or chosen for hypothesis testing.

When group is manually defined, use paste(x,y,z) or interaction(x,y,z) form; "x", "y" and "z" are the factors. There is no restrictions on the number of factors.

Value

Object of class "data.frame" containing the results.

See Also

```
P.adjust, pairwise.t.test, pair.diff, DTK.test, TukeyHSD and glht
```

Examples

mse

Mean square error

Description

Estimates the mean square error (mse)

Usage

```
mse(model)
```

Arguments

model

an object containing the results of a model.

n 21

Details

The mean square error is also known as the unexplained variance or the variance of the residuals.

Examples

n

Sample size (n)

Description

Gives n without NA's

Usage

n(x)

Arguments

х

Vector (numeric or character)

Examples

```
x= rep(c(rnorm(30,20,5),NA),3)
n(x)
```

norm.test

Normality tests

Description

 $\label{liliefors} Lilliefors (Kolmogorov-Smirnov), Shapiro-Francia, Shapiro-Wilk, D'Agostino Skewness, Anscombe-Glynn Kurtosis and D'Agostino-Pearson normality tests.$

Usage

```
## Default S3 method:
  plot(norm.test(x, title=NULL, type=c("G1","b1","mc")))
```

22 norm.test

Arguments

title

x one of:

• a numeric vector

 an object of class c('lm', 'aov', 'glm', 'gam') resulting from a calls to c(lm,aov,glm,gam)

the title at the top of the results. Default is "Normality Tests".

sk type of skewness used in D'Agostino skewness test. Can be "G1", "b1" or

"mc". Read details.

type type of residuals which should be used. See details.

Details

This function can be used on objects belonging to c('lm', 'aov', 'glm', 'gam') classes. For example, class(aov.model) gives "aov" "lm" and class(glm.model) gives "glm" "lm". The type of residuals can be defined. It generally includes c("working", "response", "deviance", "pearson", "partial").

D'Agostino-Pearson's test is more appropriate for analysing a vector with duplicate values in it. The more there are duplicate values, the more Shapiro-Wilk will be far from correctly testing the H0 hypothesis.

Given samples from a population, the equation for the sample skewness g_1 is a biased estimator of the population skewness. The use of G_1 or b_1 is advisable. For large samples, the various skewness estimates yield similar results. For small normal distributed samples, b_1 is less biased than G_1 . However, for small non-normal distributed samples, G_1 is less biased than b_1 . These two skewness estimate can be sensitive to outliers in the data (contaminated data). Therefore, the medcouple mc is also an option in type. It has a good performance on uncontaminated data and is robust on contaminated data. For more information on medcouple, please read references in mc{robustbase}.

• Typical definition used in many older textbooks:

$$g_1 = \frac{m_3}{m_2^{3/2}}$$

where m_3 is the sample third central moment, and m_2 is the sample variance.

• Definition used in SAS and SPSS:

$$G_1 = g_1 \frac{k_3}{k_2^{3/2}} = g_1 \frac{\sqrt{n(n-1)}}{n-2}$$

where k_3 is the unique symmetric unbiased estimator of the third cumulant and k_2 is the symmetric unbiased estimator of the second cumulant.

• Definition used in MINITAB and BMDP:

$$b_1 = \frac{m_3}{s^3} = g_1 \left(\frac{n-1}{n}\right)^{3/2}$$

More will be added to this section especially for Anscombe-Glynn Kurtosis test.

P.adjust 23

Value

An S4 object of class 'norm' containing the following components:

the statistics for each analysis

p.value estimated p-values based on the statistics

data original data (data.frame)

data.name names of the object called

title title for the result

References

D. N. Joanes and C. A. Gill (1998), Comparing measures of sample skewness and kurtosis. *The Statistician*, **47**, 183–189.

- G. Brys, M. Hubert and A. Struyf (2003), A Comparison of Some NewMeasures of Skewness. in *Developments in Robust Statistics* **ICORS 2001**, eds. R. Dutter, P. Filzmoser, U. Gather, and P.J. Rousseeuw, Heidelberg: Springer-Verlag, 98–113
- G. Brys, M. Hubert and A. Struyf (2004), A Robust Measure of Skewness; JCGS 13 (4), 996–1017.

Examples

```
x <- rnorm(300, 50, 10)
y <- 5*(x +10*(rnorm(300,1,2)))

norm.test(x)  ## mc skewness
norm.test(x, type="G1") ## G1 skewness
norm.test(x, type="b1") ## b1 skewness

mod <- lm(y~x)
norm.test(mod)</pre>
```

P.adjust

 $Adjust\ P-values\ for\ Multiple\ Comparisons$

Description

Given a set of p-values, returns p-values adjusted using one of several methods. This is a modified version of p.ajust from stats. It now includes "sidak" correction.

Usage

Arguments

p vector of p-values (possibly with NAs).

method correction method

n number of pvalues considered for correction; only set this (to non-default)

when you know what you are doing! See details

Details

The adjustment methods include the Bonferroni correction ("bonferroni") in which the p-values are multiplied by the number of comparisons. Less conservative corrections are also included by Holm (1979) ("holm"), Hochberg (1988) ("hochberg"), Hommel (1988) ("hommel"), Benjamini & Hochberg (1995) ("BH"), and Benjamini & Yekutieli (2001) ("BY"), respectively. A pass-through option ("none") is also included. The P.adjust.methods vector containes the set of correction methods for the benefit of methods that need to have the method as an option and pass it on to P.adjust.

The first five methods are designed to give strong control of the family wise error rate. There seems no reason to use the unmodified Bonferroni correction because it is dominated by Holm's method, which is also valid under arbitrary assumptions.

Hochberg's and Hommel's methods are valid when the hypothesis tests are independent or when they are non-negatively associated (Sarkar, 1998; Sarkar and Chang, 1997). Hommel's method is more powerful than Hochberg's, but the difference is usually small and the Hochberg p-values are faster to compute.

The "BH" and "BY" method of Benjamini, Hochberg, and Yekutieli control the false discovery rate, the expected proportion of false discoveries amongst the rejected hypotheses. The false discovery rate is a less stringent condition than the family wise error rate, so these methods are more powerful than the others.

When making multiple comparisons, n can be used to identify the number of comparisons that are actually used. Correction is then done on the full output matrix when only some of the comparisons are meaningfull or chosen for hypothesis testing. This can be done with the "bonferroni" and "sidak" correction. If other methods are used, exclude the unwanted p.values before applying correction. Unless you know what you are doing, **DO NOT** modify n if all comparisons are used. Most of the time n should be equal to length(p).

Note that you can set n larger than length(p) which means the unobserved p-values are assumed to be greater than all the observed p for "bonferroni" and "holm" methods and equal to 1 for the other methods.

Value

A vector of corrected p-values (same length as p).

References

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series* B, **57**, 289–300.

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* **29**, 1165–1188.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.

P.adjust 25

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, **75**, 383–386.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, **75**, 800–803.

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, **46**, 561–576. (An excellent review of the area.)

Sarkar, S. (1998). Some probability inequalities for ordered MTP2 random variables: a proof of Simes conjecture. *Annals of Statistics*, **26**, 494–504.

Sarkar, S., and Chang, C. K. (1997). Simes' method for multiple hypothesis testing with positively dependent test statistics. *Journal of the American Statistical Association*, **92**, 1601–1608.

Wright, S. P. (1992). Adjusted P-values for simultaneous inference. *Biometrics*, **48**, 1005–1013. (Explains the adjusted P-value approach.)

See Also

```
pairwise.t.test, mc.long, DTK.test, TukeyHSD and glht
```

```
require(graphics)
set.seed(123)
x \leftarrow rnorm(50, mean=c(rep(0,25), rep(3,25)))
p <- 2*pnorm( sort(-abs(x)))</pre>
round(p, 3)
round(P.adjust(p), 3)
round(P.adjust(p,"BH"), 3)
## or all of them at once (dropping the "fdr" alias):
P.adjust.M <- P.adjust.methods[P.adjust.methods != "fdr"]
p.adj <- sapply(P.adjust.M, function(meth) P.adjust(p, meth))</pre>
round(p.adj, 3)
## or a bit nicer:
noquote(apply(p.adj, 2, format.pval, digits = 3))
## and a graphic:
matplot(p, p.adj, ylab="P.adjust(p, meth)", type = "l", asp=1, lty=1:6,
        main = "P-value adjustments")
legend(.7,.6, P.adjust.M, col=1:6, lty=1:6)
## Can work with NA's:
pN \leftarrow p; iN \leftarrow c(46,47); pN[iN] \leftarrow NA
pN.a <- sapply(P.adjust.M, function(meth) P.adjust(pN, meth))
## The smallest 20 P-values all affected by the NA's :
round((pN.a / p.adj)[1:20, ] , 4)
```

26 pair.diff

pack.list

List of installed packages

Description

Create a text file containing the list of the packages currently installed in R.

Usage

```
pack.list(n.names=7)
```

Arguments

n.names

Number of package names to put per line of the output text file.

Author(s)

Benoit Bruneau

Examples

```
pack.list()
pack.list(5)
```

pair.diff

Mean differences matrix and their associated standard Errors

Description

Creates two lower triangle matrix: The mean differences and their standard error.

Usage

```
pair.diff(y, ...)
## S3 method for class 'formula'
pair.diff(y, data=NULL ...)
## S3 method for class 'lm'
pair.diff( y, ...)
## Default S3 method:
pair.diff( y, group, data=NULL, ...)
```

Arguments

| У | response variable for the default method, or lm or formula object. If y is a linear-model object or a formula, the variables on the right-hand-side of the model must all be factors and must be completely crossed. |
|-------|--|
| group | for the default method, factor (concatenated factor when multiple factors). See details. |
| data | ${\tt data.frame}$ where the dependant variable and the factor(s) are. |
| | additional arguments to pass to mean and/or sd. |

performance 27

Details

When group is manually defined, use paste(x,y,z) or interaction(x,y,z) form where "x", "y" and "z" are the factors. There is no restrictions on the number of factors.

This function can be usefull with pairwise.t.test since the matrix created are of the same format.

Value

Object of class "list" containing two matrices:

diff.m Mean differences half matrix

diff.se Standard error associated with the mean differences half matrix

See Also

Is included in mc.long for the long format of the results.

Examples

performance

performance

Usage

```
performance(expr, samples = 1, gcFirst = TRUE)
```

Arguments

```
expr
samples
gcFirst
```

```
# will soon be available
```

28 plot.ypr

Usage

Arguments

object an object of class 'glm' resulting from a call to glm.

se.pred logical; if TRUE, SE is plotted.

leg logical; if TRUE, a legend containing logistic equation and estimated val-

ues for the variables is plotted.

ref logical; if TRUE, reference lines for L90, L50 and L10 are plotted.

range.x the range used to define xlim in the plot. Read 'details'.

Details

More to come.

Author(s)

Benoit Bruneau

| plot.ypr | Standard Yield per Recruit plot. |
|----------|----------------------------------|
| | |

Description

Yield per Recruit and Spawning Stock Biomass per Recruit are plotted with standard reference points.

Usage

Arguments

```
an object of class "ypr" resulting from a call to ypr.1
object
                 main title for the graph
main
                 a label for the YPR y axis
ylab.ypr
ylab.ssb
                 a label for the SSB/R y axis
xlab
                 a label for the YPR x axis.
                 the color of the the color of the YPR line.
col.ypr
                 the color of the the color of the SSB/R line.
col.ssb
                 logical; if TRUE, standard reference points are added to the plot.
ref
                 logical; if TRUE, a legend is added in the 'topright' corner of the plot.
legend
```

QQplot 29

Details

More to come.

See Also

```
ypr.1
```

Examples

QQplot

QQplot

Usage

```
QQplot(dat, quant=TRUE,cex.q=2,norm=T, ...)
```

Arguments

dat one of:

- ullet a numeric vector
- an object of class c('norm', 'lm', 'aov', 'glm', 'gam') resulting from a calls to c(norm.test,lm,aov,glm,gam)

quant logical; T for adding quantiles 75, 50 (median) and 25.

cex.q numeric vector giving the amount by which plotting symbols should be

magnified relative to the default

norm logical; T adds a line to a normal quantile-quantile plot.

... additional arguments to be passed (see par, qqnorm)

```
x=rnorm(50)
QQplot(x)
norm.x=norm.test(x)
QQplot(norm.x)
```

30 reject.z

r.colors

Pie charts of all R character colors

Description

Creates a pdf file with pie charts of all the 657 basic character colors of R

Usage

```
r.colors(file)
```

Arguments

file

the directory in which the pdf file will be created

Details

Define the directory in which the file should saved by writing file="C:/temp" for example. If file is not defined, it will be saved in "C:/" on windows and in "home" on Mac.

Value

None

Examples

```
r.colors()
```

reject.z

reject z

${\bf Usage}$

```
reject.z(x, index = NULL, threshold = 2)
```

Arguments

x

index

threshold

Examples

will soon be available

replace.z 31

```
replace z
 replace.z
Usage
   replace.z(x, index = NULL, threshold = 2)
Arguments
   X
   index
   threshold
Examples
   # will soon be available
 resid.ortho
                       Orthogonal\ residuals
Usage
   xxx(data,,)
Arguments
   data
Author(s)
   Benoit Bruneau
```

Description

rivard

This function applies Rivard equations to mid-year weight at age data to adjust values to Jan-1 basis.

Rivard Weights Calculation

Usage

```
rivard(pds, pred=FALSE, K=2, plus.gr=FALSE)
```

Arguments

data

32 rollmin

Details

More to come. Will be adding interpolation for spawning season.

Examples

```
x=rnorm(30,800,10)
rivard(data.frame("2000"=x,"2001"=x*1.2, "2002"=x*0.8,"2003"=x*0.5))
```

rm.levels

rm factor levels

Usage

```
rm.levels(factor)
```

Arguments

factor

Examples

will soon be available

rollmin

rollmin

Usage

Arguments

x k na.pad align

Examples

. . .

```
# will soon be available
```

roundup 33

 $roundup \hspace{3cm} roundup$

Description

The "conventional" rounding of 5 to the higher value

Usage

```
roundup(x, numdigits = 0)
```

Arguments

x numeric vector.

digits integer indicating the number of decimal places to be used.

Details

Rounds a 5 to the next value. Therefore roundup(2.5) is 3. This can be usefull when the rounded values are to be presented in a document (eg. table, graph,...).

When rounded values are used in other calculations, **round** should be used since it follows the IEC 60559 standard.

Value

numeric vector.

See Also

round

Examples

round(2.5)
roundup(2.5)

runmax runmax

Usage

```
runmax(x, window)
```

Arguments

х

window

```
# will soon be available
```

s.an

runmean runmeanUsage runmean(x, window) Arguments window Examples # will soon be available runmin runmin ${\bf Usage}$ runmin(x, window) Arguments х window Examples # will soon be available $Simulations \ for \ YPR \ model$ s.an Description Not ready yet. Use for loops for now. Usage xxx(data , , ,) Arguments data

Author(s)

Benoit Bruneau

show.North 35

se Standard Error

Usage

```
se(x, na.rm=T)
```

Arguments

x an R object (vector, matrix,...)

na.rm a logical value indicating whether NA values should be stripped before

the computation proceeds

Details

The standard error of the mean is defined as:

$$SE = \frac{sd}{\sqrt{n}}$$

where sd is the standard deviation of the sample and n is the sample size.

Examples

x=rnorm(50)
se(x)

show.North

North arrow for a map

Description

Draws North arrow on a map

Usage

Arguments

pos Position of the arrow. Default is 'topright'. See details.

arrow.col Arrow color.

arrow.fill Color inside the head of the arrow. NA for no color.

arrow.lwd Line width of the arrow.

N.cex Character size for 'N'.

N.family Font family of 'N'.

36 sort.vdf

Details

The position of the north arrow is defined by pos and can either be numeric or character.

If pos is a numeric vector, it is a vector of the form c(x,y) where x and y are fractions of the plotting region. If x and y are not in [0,1], then the north arrow is drawn outside the bounds of the plotting region and a warning message is given.

If pos is a character vector, it should be a single keyword from:

```
• c('topright', 'topleft', 'bottomright', 'bottomleft')
```

Examples

```
plot(1)
show.North()
show.North(c(0.8,0.9))
show.North(c(1.01,0.9)) ### gives a warning
```

sort.vdf

Sort Data Frames and Vectors

Description

Single function enabling data.frame and vector sorting

Usage

```
sort.vdf(x, by, increasing=TRUE)
```

Arguments

x data.frame or vector

by A one-sided formula using + for ascending and - for descending. Sorting

is left to right in the formula. This is for data.frame only.

increasing logical. Should the sort be increasing (TRUE) or decreasing (FALSE)? This

is for sorting vectors only.

Details

See example.

Author(s)

Kevin Wright and modified by Benoit Bruneau

summary.ypr 37

Examples

summary.ypr

Summarizing the results of YPR models.

Description

Summary for an object of class "ypr".

Usage

```
## S4 method for signature 'ypr'
summary(object)
```

Arguments

object an object of class "ypr" resulting from a call to ypr.1.

Examples

ttest.perm

Student's t-tests by Permutation

Description

Performs two sample t-tests or paired t-test by use of permutation

Usage

38 ttest.perm

Arguments

vec1, vec2 two numeric vectors used for Student's t-test analysis number of permutations (default = 999) nperm one of the following: "two.sided", "less" or "greater". alternative a logical variable indicating whether to treat the two variances as being var.equal equal (TRUE) or not (FALSE). silent a logical variable indicating whether calculation results are printed (FALSE) to the R console or not (TRUE).

one of the following: "i" for independant samples or "p" for paired samples. type

exact a logical variable indicating whether to perform the exact test (TRUE) or

not (FALSE).

Details

The permutational t-test does not require normality of the distributions of each variable. It is also quite robust to heteroscedasticity.

Use exact=TRUE to perform two sample t-test on all the possible combination. This option can only be used when the sum of the sample sizes $(n_1 + n_2)$ is smaller than 20. It is recommended to use this option when sample sizes are small. It is not implemented yet in the paired t-test.

nperm can not be higher than the maximum number of combination possible (n_{comb}) .

```
n_{comb} = N!/(n_1!n_2!)
     where n_comb is the number of possible combinations, N! is factorial (n_1 + n_2), n_1!
     is factorial(n(vec1)) and n_2! is factorial(n(vec2)).
```

There is more to come in this section.

Value

t.ref reference value of the t-statistic p.param parametric p-value permutational p-value p.perm number of permutations nperm list of the t statistics (only for independent sample ttest), starting with perm.t

the reference value, followed by all values obtained under permutations.

```
x \leftarrow rnorm(50,0,1)
y \leftarrow runif(50,0,1)*x
toto = ttest.perm(x, y) ##independant samples ttest
```

unload 39

unload

 $Unload\ packages$

Description

Unloads one or multiple packages.

Usage

unload(pack)

Arguments

pack

Character vector specifying which packages to unload.

Author(s)

Benoit Bruneau

Examples

```
library(mgcv)
search()
unload(mgcv)
search()
```

week.1

week.1

Description

Week of the year starting on the first of January (01-53)

Usage

week.1(x)

Arguments

X

Author(s)

Denis Chabot

Examples

will soon be available

week.num

week.num

Description

Week of the year as decimal number (00-53) using Sunday or Monday as the first day 1 of the week (and typically with the first Sunday of the year as day 1 of week 1).

Usage

```
week.num(x, day=c("sunday", "monday"))
```

Arguments

```
x A vector of dates.
```

day Either "sunday" or "monday". Default is "sunday".

Details

Argument day indicates if the week starts on "sunday" or "monday".

Examples

ypr.1

Length Based Yield Per Recruit

Description

Length based Yield Per Recruit model is define by fishery selectivity and life history parameters related to length.

Usage

```
ypr.1(LW, vonB, 1.start, last.age, age.step=1, Fsel.type,
     F.max=2,F.incr.YPR=0.0001, M.1, M=0.2, f.MSP=0.4,
     F.f=0, M.f=0.5, riv.calc=FALSE)
```

Arguments

LW one of:

• a vector containing $c(\alpha, \beta)$ from length-weight curve. See 'Details'.

• an object of class 'nls' in which α and β were estimated. See 'Details'.

vonB one of:

a vector containing c(Linf,K) from von Bertalanffy grotwh curve.
 See 'Details'.

an object of class 'glm' in which eqnLinf and eqnK were estimated.
 See 'Details'.

1.start length at the starting age

last age to be considered in the model

age.step steps used to generate ages. Default is 1.

Fsel.type one of:

• a list containing the type of fishery selectivity and the values needed for the function related to the type. See 'Details'.

• an object of class 'glm' in which α and β were estimated by a logistic regression. See 'Details'.

F.max maximum value of instantaneous rate of fishing mortality (F). Default is 2.

F.incr.YPR increment for generating the F values to be used for YPR calculation. Default is 0.0001.

M.1 one of:

• a list containing the type of maturity at length definition and the values needed for the function related to the type. See 'Details'.

• an object of class 'glm' in which α and β were estimated by a logistic regression. See 'Details'.

M instantaneous rate of natural mortality (M). Default is 0.2.

f.MSP reference point defined as the fraction of maximum spawning potential.

Default is 0.4.

F.f fraction of instantaneous rate of fishing mortality (F) before spawning.

M.f fraction of instantaneous rate of natural mortality (M) before spawning.

riv.calc a logical value indicating whether to use Rivard weights calculation (TRUE) or not. Default is FALSE.

Details

Length-Weight relationship can be provided either by indicating $c(\alpha, \beta)$ values in a vector or by directly using an object of class 'nls' or 'lm'. If α and β are estimated by lm, log(x, base=exp(1)) transformation should be applied to the data prior to fitting the linear model. If an object resulting from nls is used, variables should be named alpha and beta using the following equation:

$$W = \alpha L^{\beta}$$

where W is weight, L is length, α is the elevation of the curve, and β is the steepness of

the curve. Both α and β are coefficients estimated by the regression.

Von Bartalanffy growth equation parameters can be provided either by indicating c(Linf, K) values in a vector or by directly using an object of class 'nls'. If an object resulting from nls is used, variables should be named Linf and K. As for t_0 , any name may be used since only L_{∞} and K are used in this length-based YPR model. The equation used in the nls for estimating L_{∞} and K should be the following one:

$$L_t = L_{\infty} \left(1 - e^{-K(t - t_0)} \right)$$

where L_t is length-at-age t, L_{∞} is the asymptotic average maximum length, K is a growth rate coefficient determinant of how quick the maximum is attained, and t_0 is the hypothetical age at length zero.

As stated above, since this length-based YPR model uses relative age, $t - t_0$ becomes a relative age (a). The Von Bartalanffy growth equation used in this length-based YPR model is defined as:

$$L_a = L_{\infty} \left(1 - e^{-Ka} \right) + L_s e^{-Ka}$$

where L_a is length at a relative age a and L_s is length at relative age zero.

The fishery selectivity and maturity at length components of the model can be defined as one of c("full", "ramp", "logistic") equations. The proper way to specify which equation to use is by the construct of a list where the first element is the name of one of the three types of equation. The following elements of the list are specific to the type of equation:

- full: element [[2]] is the length at which full maturity is achieved.
- ramp: element [[2]] is the maximum length at which maturity is null and element [[3]] is the minimum length at which maturity is fully achieved.
- logistic elements [[2]] and [[3]] are respectively α and β components of a logistic curve.

Alternatively, an object of class 'glm' can directly be used for the **fishery selectivity** and **maturity at length** components. The Generalized Linear Model should have the option family set to either binomial or quasibinomial keeping link function to the default (*i.e.* "logit"). Estimated coefficients are use as follow:

$$y = \frac{1}{1 + e^{-(\alpha + \beta x)}}$$

Reference points used for result output are defined as:

- **F.zero:** F level when there is no fishing (F=0).
- F.01: F level where the slope of yield curve is 10% of the slope at F.zero.
- **F.xx:** F level where the MSP is at the level defined by **f.**MSP option. Default is 40% (0.4).

• F.max: F level where yield is maximum.

More to come.

Value

ypr.1 returns an object of class(S4) "ypr". The functions summary and plot are used to respectively obtain a summary and a standard plot of the results.

An object of class "ypr" has the the following slots:

parms the list of parameters used in the model.

base a data.frame containing the starting values:

• relative age classes

length at ageweight at age

refs a data.frame containing values predicted by the model for the four ref-

erence points. See details.

YPR a data.frame containing the results for all partial Fs.

Note that to have access to each slot of an "ypr" object, one must use "@" instead of "\$".

Author(s)

Benoit Bruneau

See Also

```
plot.logit
```

Index

| *Topic design | leveneTest, 16 |
|-------------------------|---------------------------------|
| lsmean, 17 | lib.code, 16 |
| *Topic ttest | list, 10 |
| ttest.perm, 37 | lm, 11, 13, 22, 29 |
| oceso.perm, or | lsmean, 17 |
| aov, 11, 22, 29 | ismean, ir |
| arrows, 8 | make.z, 19 |
| att.strp, 2 | mc, 22 |
| | mc.long, 19, 25, 27 |
| bmisc, 4 | mean, 26 |
| | mse, 20 |
| ${\tt ceiling.lg,5}$ | mtext, 10 |
| clean, 5 | mocko, 10 |
| corr.perm, 6 | n, 21, <i>38</i> |
| cv, 7 | NA, 24 |
| | nclass.FD, 12 |
| data.frame, 15, 19 | nclass.scott, 12 |
| day, 7 | nclass.Sturges, 12 |
| DTK.test, 20 , 25 | nls, 42 |
| | norm.test, 11, 21, 29 |
| Errbar, 8 | 101111.0000, 11, 21, 20 |
| expression, 10 | P.adjust, 19, 20, 23 |
| £+1 00 | p.ajust, 23 |
| factorial, 38 | pack.list, 26 |
| fct, 9 | pair.diff, 20, 26 |
| for, 34 | pairwise.t.test, 19, 20, 25, 27 |
| format.hms, 9 | par, 8, 10, 12, 29 |
| formula, $19, 26$ | paste, 20 |
| com 11 00 00 | performance, 27 |
| gam, 11, 22, 29 | plot.logit, 28, 43 |
| gam.Check, 10 | plot.ypr, 28 |
| gam.check, 10 | predict, 18 |
| get.partial.etas, 11 | prodres, re |
| glht, 20, 25 | qqnorm, 29 |
| glm, 11, 13, 22, 28, 29 | QQplot, 29 |
| hist, 12 | , |
| histplot, 10, 11 | $	exttt{r.colors}, 30$ |
| піверіос, 10, 11 | reject.z, 30 |
| interaction, 20 | replace.z, 31 |
| inv.pred, 12 | resid.ortho, 31 |
| is.even, 13, 14 | rivard, 31 |
| is.odd, 13, 14 | rm.levels, 32 |
| 20.000, 20, 21 | robustbase, 22 |
| last, 14 | rollmin, 32 |
| lev, 15 | round, 33 |
| · | * |

INDEX 45

```
roundup, 13, 14, 33
rug, 12
\mathtt{runmax},\, 33
runmean, 34
\mathtt{runmin},\, {\color{red} 34}
s.an, 34
sd, 26
{\tt se},\, {\color{red}35}
\mathtt{show.North},\, \color{red} 35
\mathtt{sort.vdf},\, 36
stats, 23
summary.ypr, 37
t.test, 19
{\tt ttest.perm},\, {\tt 37}
TukeyHSD, 20, 25
unload, 39
{\tt week.1},\, 39
{\tt week.num},\,40
ypr.1, 28, 29, 37, 40
```