

Package ‘rcompanion’

February 19, 2018

Version 1.11.3

Date 2018-02-19

Title Functions to Support Extension Education Program Evaluation

Description Functions and datasets to support “Summary and Analysis of Extension Program Evaluation in R” and “An R Companion for the Handbook of Biological Statistics”.
Vignettes are available at <<http://rcompanion.org>>.

URL <http://rcompanion.org/>

Depends R (>= 3.3.0)

Imports stats (>= 3.3.0), graphics (>= 3.3.0), boot (>= 1.3-18), DescTools (>= 0.99.17), multcompView (>= 0.1-7), plyr (>= 1.8.4), EMT (>= 1.1), ordinal (>= 2015.6-28), coin (>= 1.1-2), lattice (>= 0.20-33), RVAideMemoire (>= 0.9-60), WRS2 (>= 0.9-1), lmtest (>= 0.9-34), BSDA (>= 1.01), nortest (>= 1.0-4), hermite (>= 1.1.1)

Suggests FSA (>= 0.8.9)

Author Salvatore Mangiafico [aut, cre]

Maintainer Salvatore Mangiafico <mangiafico@njaes.rutgers.edu>

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2018-02-19 14:13:02 UTC

R topics documented:

rcompanion-package	3
accuracy	4

Anderson	6
AndersonBias	6
AndersonRainBarrel	7
AndersonRainGarden	7
BobBelcher	8
Breakfast	8
BrendonSmall	9
BullyHill	9
Catbus	10
cateNelson	10
cldList	12
cohenW	14
compareGLM	15
compareLM	16
cramerV	18
cramerVFit	19
epsilonSquared	20
freemanTheta	22
fullPTable	23
groupwiseCMH	24
groupwiseGeometric	25
groupwiseHuber	27
groupwiseMean	28
groupwiseMedian	30
groupwisePercentile	32
groupwiseSum	34
HayleySmith	35
Monarchs	36
nagelkerke	36
nagelkerkeHermite	38
nominalSymmetryTest	39
Nurseries	41
pairwiseDifferences	41
pairwiseMcnemar	43
pairwiseMedianMatrix	45
pairwiseMedianTest	46
pairwiseModelAnova	48
pairwiseNominalIndependence	49
pairwiseNominalMatrix	50
pairwiseOrdinalIndependence	51
pairwiseOrdinalMatrix	53
pairwiseOrdinalPairedMatrix	54
pairwiseOrdinalPairedTest	56
pairwiseOrdinalTest	57
pairwisePercentileTest	59
pairwisePermutationMatrix	61
pairwisePermutationSymmetry	62
pairwisePermutationSymmetryMatrix	64

pairwisePermutationTest	65
pairwiseRobustMatrix	67
pairwiseRobustTest	68
pairwiseSignMatrix	70
pairwiseSignTest	71
percentileTest	73
plotDensityHistogram	75
plotNormalDensity	76
plotNormalHistogram	77
plotPredy	79
Pooh	81
PoohPiglet	81
Religion	82
scheirerRayHare	82
transformTukey	83
Index	85

rcompanion-package	<i>rcompanion: Functions to support extension education program evaluation</i>
--------------------	--

Description

This package provides custom functions for working through examples and analyses from "Summary and Analysis of Extension Education Program Evaluation in R" and "An R Companion for the Handbook of Biological Statistics".

Useful functions

The function [nagelkerke](#) provides pseudo R-squared values for a variety of model types, as well as a likelihood ratio test for the model as a whole. An additional function, [nagelkerkeHermite](#), is provided for models fit with the `hermite` package.

There are several functions that provide summary statistics for grouped data. These function titles tend to start with "groupwise". They provide means, medians, geometric means, and Huber M-estimators for groups, along with confidence intervals by traditional methods and bootstrap.

Function titles starting with "pairwise" conduct pairwise tests among groups as a post-hoc analysis for omnibus tests. At the time of writing, these tests are Mood's median test, sign test (for omnibus Friedman test), permutation test, robust anova, and ordinal regression. The output is a table of comparisons and p-values, or a matrix of p-values that can be parsed into a compact letter display.

There are also functions that are useful for comparing models. [compareLM](#), [compareGLM](#), and [pairwiseModelAnova](#). These use goodness-of-fit measures like AIC, BIC, and BICc, or likelihood ratio tests.

There are a few useful plotting functions, including [plotNormalHistogram](#) that plots a histogram of values and overlays a normal curve, and [plotPredy](#) which plots of line for predicted values for a bivariate model. Other plotting functions include producing density plots.

Functions for nominal data include post-hoc tests for Cochran-Mantel-Haenszel test ([groupwiseCMH](#)), for McNemar-Bowker test ([pairwiseMcNemar](#)), and for tests of association like Chi-square, Fisher exact, and G-test ([pairwiseNominalIndependence](#)).

A function close to my heart is ([cateNelson](#)), which performs Cate-Nelson analysis for bivariate data.

Vignettes and examples

The functions in this package are used in "Extension Education Program Evaluation in R" which is available at <http://rcompanion.org/handbook/> and "An R Companion for the Handbook of Biological Statistics" which is available at <http://rcompanion.org/rcompanion/>.

The documentation for each function includes an example as well.

accuracy	<i>Minimum maximum accuracy, mean absolute percent error, root mean square error, and Efron's pseudo r-squared</i>
----------	--

Description

Produces a table of fit statistics for multiple models.

Usage

```
accuracy(fits, plotit = TRUE, digits = 3, ...)
```

Arguments

<code>fits</code>	A series of model object names. Must be a list.
<code>plotit</code>	If TRUE, produces plots of the predicted values vs. the actual values for each model.
<code>digits</code>	The number of significant digits in the output.
<code>...</code>	Other arguments passed to <code>plot</code> .

Details

Produces a table of fit statistics for multiple models: minimum maximum accuracy, mean absolute percentage error, root mean square error, normalized root mean square error, accuracy based on normalized root mean square error, and Efron's pseudo r-squared.

For minimum maximum accuracy, larger indicates a better fit, and a perfect fit is equal to 1.

For mean absolute error (MAE), smaller indicates a better fit, and a perfect fit is equal to 0. It has the same units as the dependent variable.

For mean absolute percent error (MAPE), smaller indicates a better fit, and a perfect fit is equal to 0.

Root mean square error (RMSE) has the same units as the predicted values.

Normalized root mean square error (NRMSE) is RMSE divided by the mean or the median of the values of the dependent variable.

NRMSE accuracy values are calculated as 1 minus NRMSE. Larger indicates a better fit, and a perfect fit is equal to 1.

Efron's pseudo r-squared is calculated as 1 minus the residual sum of squares divided by the total sum of squares. For linear models (lm model objects), Efron's pseudo r-squared will be equal to r-squared. For other models, it should not be interpreted as r-squared, but can still be useful as a relative measure.

Model objects currently supported: lm, glm, nls, betareg, gls, lme, lmer, lmerTest, rq, loess, gam, glm.nb, glmRob.

Value

A list of two objects: The series of model calls, and a data frame of statistics for each model.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/G_14.html

See Also

[compareLM](#), [compareGLM](#), [nagelkerke](#)

Examples

```
data(BrendonSmall)
BrendonSmall$Calories = as.numeric(BrendonSmall$Calories)
BrendonSmall$Calories2 = BrendonSmall$Calories ^ 2
model.1 = lm(Sodium ~ Calories, data = BrendonSmall)
model.2 = lm(Sodium ~ Calories + Calories2, data = BrendonSmall)
model.3 = glm(Sodium ~ Calories, data = BrendonSmall, family="Gamma")
quadplat = function(x, a, b, clx) {
  ifelse(x < clx, a + b * x + (-0.5*b/clx) * x * x,
         a + b * clx + (-0.5*b/clx) * clx * clx)}
model.4 = nls(Sodium ~ quadplat(Calories, a, b, clx),
              data = BrendonSmall,
              start = list(a=519, b=0.359, clx = 2300))
accuracy(list(model.1, model.2, model.3, model.4), plotit=FALSE)

### Perfect and poor model fits
X = c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)
Y = c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)
Z = c(1, 12, 13, 6, 10, 13, 4, 3, 5, 6, 10, 14)
perfect = lm(Y ~ X)
poor    = lm(Z ~ X)
accuracy(list(perfect, poor), plotit=FALSE)
```

Anderson

Hypothetical data for Alexander Anderson

Description

A matrix of counts for students passing or failing a pesticide training course across four counties. Hypothetical data.

Usage

Anderson

Format

An object of class `matrix` with 4 rows and 2 columns.

Source

http://rcompanion.org/handbook/H_04.html

AndersonBias

Hypothetical data for Alexander Anderson with sex bias

Description

A data frame of counts for students passing or failing a pesticide training course across four counties, with sex of students. Hypothetical data.

Usage

AndersonBias

Format

An object of class `data.frame` with 16 rows and 4 columns.

Source

http://rcompanion.org/handbook/H_06.html

AndersonRainBarrel *Hypothetical data for Alexander Anderson on rain barrel installation*

Description

A matrix of paired counts for students planning to install rain barrels before and after a class. Hypothetical data.

Usage

```
AndersonRainBarrel
```

Format

An object of class `matrix` with 2 rows and 2 columns.

Source

http://rcompanion.org/handbook/H_05.html

AndersonRainGarden *Hypothetical data for Alexander Anderson on rain garden installation*

Description

A matrix of paired counts for students planning to install rain gardens before and after a class. Hypothetical data.

Usage

```
AndersonRainGarden
```

Format

An object of class `matrix` with 3 rows and 3 columns.

Source

http://rcompanion.org/handbook/H_05.html

BobBelcher

Hypothetical data for ratings of instructors in unreplicated CBD

Description

A data frame of Likert responses for five instructors for each of 8 respondents. Arranged in unreplicated complete block design. Hypothetical data.

Usage

BobBelcher

Format

An object of class `data.frame` with 40 rows and 3 columns.

Source

http://rcompanion.org/handbook/F_10.html

Breakfast

Hypothetical data for students' breakfast habits and travel to school

Description

A two-dimensional contingency table, in which Breakfast is an ordered nominal variable, and Travel is a non-ordered nominal variable. Hypothetical data.

Usage

Breakfast

Format

An object of class `table` with 3 rows and 5 columns.

Source

http://rcompanion.org/handbook/H_09.html

`BrendonSmall`*Hypothetical data for Brendon Small and company*

Description

A data frame of the intake of calories and sodium for students in five classes. Hypothetical data.

Usage

```
BrendonSmall
```

Format

An object of class `data.frame` with 45 rows and 6 columns.

Source

http://rcompanion.org/handbook/I_10.html

`BullyHill`*Hypothetical data for proportion of students passing a certification*

Description

A data frame of counts of students passing and failing. Hypothetical data.

Usage

```
BullyHill
```

Format

An object of class `data.frame` with 12 rows and 5 columns.

Source

http://rcompanion.org/handbook/J_02.html

Catbus	<i>Hypothetical data for Catbus and company</i>
--------	---

Description

A data frame of the number of steps taken by students in three classes. Hypothetical data.

Usage

Catbus

Format

An object of class `data.frame` with 26 rows and 5 columns.

Source

http://rcompanion.org/handbook/C_03.html

cateNelson	<i>Cate-Nelson models for bivariate data</i>
------------	--

Description

Produces critical-x and critical-y values for bivariate data according to a Cate-Nelson analysis.

Usage

```
cateNelson(x, y, plotit = TRUE, hollow = TRUE, xlab = "X", ylab = "Y",
  trend = "positive", clx = 1, cly = 1, xthreshold = 0.1,
  ythreshold = 0.1, progress = TRUE, verbose = TRUE, listout = FALSE)
```

Arguments

x	A vector of values for the x variable.
y	A vector of values for the y variable.
plotit	If TRUE, produces plots of the output.
hollow	If TRUE, uses hollow circles on the plot to indicate data not fitting the model.
xlab	The label for the x-axis.
ylab	The label for the y-axis.
trend	"positive" if the trend of y vs. x is generally positive. "negative" if negative.
clx	Indicates which of the listed critical x values should be chosen as the critical x value for the final model.

cly	Indicates which of the listed critical y values should be chosen as the critical y value for the final model.
xthreshold	Indicates the proportion of potential critical x values to display in the output. A value of 1 would display all of them.
ythreshold	Indicates the proportion of potential critical y values to display in the output. A value of 1 would display all of them.
progress	If TRUE, prints an indicator of progress as for loops progress.
verbose	If FALSE, suppresses printed output of tables.
listout	If TRUE, outputs a list of data frames instead of a single data frame. This allows a data frame of critical values and associated statistics to be extracted, for example if one would want to sort by Cramer's V.

Details

Cate-Nelson analysis divides bivariate data into two groups. For data with a positive trend, one group has a large x value associated with a large y value, and the other group has a small x value associated with a small y value. For a negative trend, a small x is associated with a large y, and so on.

The analysis is useful for bivariate data which don't conform well to linear, curvilinear, or plateau models.

This function will fail if either of the largest two or smallest two x values are identical.

Value

A data frame of statistics from the analysis: number of observations, critical level for x, sum of squares, critical value for y, the number of observations in each of the quadrants (I, II, III, IV), the number of observations that conform with the model, the proportion of observations that conform with the model, the number of observations that do not conform to the model, the proportion of observations that do not conform to the model, a p-value for the Fisher exact test for the data divided into the groups indicated by the model, and Cramer's V for the data divided into the groups indicated by the model.

Output also includes printed lists of critical values, explanation of the values in the data frame, and plots: y vs. x; sum of squares vs. critical x value; the number of observations that do not conform to the model vs. critical y value; and y vs. x with the critical values shown as lines on the plot, and the quadrants labeled.

Note

The method in this function follows *Cate, R. B., & Nelson, L.A. (1971). A simple statistical procedure for partitioning soil test correlation data into two classes. Soil Science Society of America Proceedings 35, 658-660.*

An earlier version of this function was published in *Mangiafico, S.S. 2013. Cate-Nelson Analysis for Bivariate Data Using R-project. J.of Extension 51:5, 5TOT1.*

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/rcompanion/h_02.html

Examples

```
data(Nurseries)
cateNelson(x      = Nurseries$Size,
           y      = Nurseries$Proportion,
           plotit = TRUE,
           hollow = TRUE,
           xlab   = "Nursery size in hectares",
           ylab   = "Proportion of good practices adopted",
           trend  = "positive",
           clx    = 1,
           xthreshold = 0.10,
           ythreshold = 0.15)
```

cldList

Compact letter display for lists of comparisons

Description

Produces a compact letter display (cld) from pairwise comparisons that were summarized in a table of comparisons

Usage

```
cldList(formula = NULL, data = NULL, comparison = NULL, p.value = NULL,
        threshold = 0.05, print.comp = FALSE, remove.space = TRUE,
        remove.equal = TRUE, remove.zero = TRUE, swap.colon = TRUE,
        swap.vs = FALSE, ...)
```

Arguments

formula	A formula indicating the variable holding p-values and the variable holding the comparisons. e.g. P.adj ~ Comparison.
data	The data frame to use.
comparison	A vector of text describing comparisons, with each element in a form similar to "Treat.A - Treat.B = 0". Spaces and "=" and "0" are removed by default
p.value	A vector of p-values corresponding to the comparisons in the comparison argument
threshold	The alpha value. That is, the p-value below which the comparison will be considered significant
print.comp	If TRUE, prints out a data frame of the modified text of the comparisons. Useful for debugging

<code>remove.space</code>	If TRUE, removes spaces from the text of the comparisons
<code>remove.equal</code>	If TRUE, removes "=" from the text of the comparisons
<code>remove.zero</code>	If TRUE, removes "0" from the text of the comparisons
<code>swap.colon</code>	If TRUE, swaps ":" with "-" in the text of the comparisons
<code>swap.vs</code>	If TRUE, swaps "vs" with "-" in the text of the comparisons
<code>...</code>	Additional arguments passed to <code>multcompLetters</code>

Details

The input should include either formula and data; or comparison and p.value.

This function relies upon the `multcompLetters` function in the `multcompView` package. The text for the comparisons passed to `multcompLetters` should be in the form "Treat.A-Treat.B". Currently `cldList` removes spaces, equal signs, and zeros, by default, and so can use text in the form e.g. "Treat.A - Treat.B = 0". It also changes ":" to "-", and so can use text in the form e.g. "Treat.A : Treat.B".

Value

A data frame of group names, group separation letters, and monospaced separation letters

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_08.html

Examples

```
data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                           levels=c("Pooh", "Tigger", "Piglet"))

library(FSA)
DT = dunnTest(Likert ~ Speaker,
             data=PoohPiglet,
             method="bh")

DT = DT$res
DT
cldList(P.adj ~ Comparison,
       data = DT,
       threshold = 0.05)
```

cohenW

Cohen's w (omega)

Description

Calculates Cohen's w for a table of nominal variables.

Usage

```
cohenW(x, y = NULL, p = NULL, digits = 4, ...)
```

Arguments

x	Either a two-way table or a two-way matrix. Can also be a vector of observations for one dimension of a two-way table.
y	If x is a vector, y is the vector of observations for the second dimension of a two-way table.
p	If x is a vector of observed counts, p can be given as a vector of expected probabilities, as in a chi-square goodness of fit test.
digits	The number of significant digits in the output.
...	Additional arguments passed to <code>chisq.test</code> .

Details

Cohen's w is used as a measure of association between two nominal variables, or as an effect size for a chi-square test of association. For a 2 x 2 table, the absolute value of the phi statistic is the same as Cohen's w. The value of Cohen's w is not bound by 1 on the upper end.

Value

A single statistic, Cohen's w.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_10.html

See Also

[cramerV](#) [cramerVFit](#)

Examples

```

### Example with table
data(Anderson)
fisher.test(Anderson)
cohenW(Anderson)

### Example for goodness-of-fit
### Bird foraging example, Handbook of Biological Statistics
observed = c(70, 79, 3, 4)
expected = c(0.54, 0.40, 0.05, 0.01)
chisq.test(observed, p = expected)
cohenW(observed, p = expected)

### Example with two vectors
Species = c(rep("Species1", 16), rep("Species2", 16))
Color = c(rep(c("blue", "blue", "blue", "green"),4),
          rep(c("green", "green", "green", "blue"),4))
fisher.test(Species, Color)
cohenW(Species, Color)

```

compareGLM

Compare fit statistics for glm models

Description

Produces a table of fit statistics for multiple glm models.

Usage

```
compareGLM(fits, ...)
```

Arguments

<code>fits</code>	A series of model object names, separated by commas.
<code>...</code>	Other arguments passed to <code>list</code> .

Details

Produces a table of fit statistics for multiple glm models: AIC, AICc, BIC, p-value, pseudo R-squared (McFadden, Cox and Snell, Nagelkerke).

Smaller values for AIC, AICc, and BIC indicate a better balance of goodness-of-fit of the model and the complexity of the model. The goal is to find a model that adequately explains the data without having too many terms.

BIC tends to choose models with fewer parameters relative to AIC. For comparisons with AIC, etc., to be valid, both models must have the same data, without transformations, use the same dependent variable, and be fit with the same method. They do not need to be nested.

Value

A list of two objects: The series of model calls, and a data frame of statistics for each model.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/rcompanion/e_07.html

See Also

[compareLM](#), [pairwiseModelAnova](#), [accuracy](#)

Examples

```
### Compare among logistic regression models
data(AndersonBias)
model.0 = glm(Result ~ 1, weight = Count, data = AndersonBias,
              family = binomial(link="logit"))
model.1 = glm(Result ~ County, weight = Count, data = AndersonBias,
              family = binomial(link="logit"))
model.2 = glm(Result ~ County + Sex, weight = Count, data = AndersonBias,
              family = binomial(link="logit"))
model.3 = glm(Result ~ County + Sex + County:Sex, weight = Count,
              data = AndersonBias, family = binomial(link="logit"))
compareGLM(model.0, model.1, model.2, model.3)
```

compareLM

Compare fit statistics for lm models

Description

Produces a table of fit statistics for multiple lm models.

Usage

```
compareLM(fits, ...)
```

Arguments

`fits` A series of model object names, separated by commas.
`...` Other arguments passed to `list`.

Details

Produces a table of fit statistics for multiple lm models: AIC, AICc, BIC, p-value, R-squared, and adjusted R-squared.

Smaller values for AIC, AICc, and BIC indicate a better balance of goodness-of-fit of the model and the complexity of the model. The goal is to find a model that adequately explains the data without having too many terms.

BIC tends to choose models with fewer parameters relative to AIC.

In the table, Shapiro.W and Shapiro.p are the W statistic and p-value for the Shapiro-Wilks test on the residuals of the model.

For comparisons with AIC, etc., to be valid, both models must have the same data, without transformations, use the same dependent variable, and be fit with the same method. They do not need to be nested.

Value

A list of two objects: The series of model calls, and a data frame of statistics for each model.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/I_10.html http://rcompanion.org/rcompanion/e_05.html

See Also

[compareGLM](#), [pairwiseModelAnova](#), [accuracy](#)

Examples

```
### Compare among polynomial models
data(BrendonSmall)
BrendonSmall$Calories = as.numeric(BrendonSmall$Calories)

BrendonSmall$Calories2 = BrendonSmall$Calories * BrendonSmall$Calories
BrendonSmall$Calories3 = BrendonSmall$Calories * BrendonSmall$Calories *
  BrendonSmall$Calories
BrendonSmall$Calories4 = BrendonSmall$Calories * BrendonSmall$Calories *
  BrendonSmall$Calories * BrendonSmall$Calories
model.1 = lm(Sodium ~ Calories, data = BrendonSmall)
model.2 = lm(Sodium ~ Calories + Calories2, data = BrendonSmall)
model.3 = lm(Sodium ~ Calories + Calories2 + Calories3, data = BrendonSmall)
model.4 = lm(Sodium ~ Calories + Calories2 + Calories3 + Calories4,
  data = BrendonSmall)
compareLM(model.1, model.2, model.3, model.4)
```

`cramerV`*Cramer's V (phi)*

Description

Calculates Cramer's V for a table of nominal variables.

Usage

```
cramerV(x, y = NULL, digits = 4, bias.correct = FALSE, ...)
```

Arguments

<code>x</code>	Either a two-way table or a two-way matrix. Can also be a vector of observations for one dimension of a two-way table.
<code>y</code>	If <code>x</code> is a vector, <code>y</code> is the vector of observations for the second dimension of a two-way table.
<code>digits</code>	The number of significant digits in the output.
<code>bias.correct</code>	If TRUE, a bias correction is applied.
<code>...</code>	Additional arguments passed to <code>chisq.test</code> .

Details

Cramer's V is used as a measure of association between two nominal variables, or as an effect size for a chi-square test of association. For a 2 x 2 table, the absolute value of the phi statistic is the same as Cramer's V.

Value

A single statistic, Cramer's V.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_10.html

See Also

[cohenW](#)

Examples

```
### Example with table
data(Anderson)
fisher.test(Anderson)
cramerV(Anderson)

### Example with two vectors
Species = c(rep("Species1", 16), rep("Species2", 16))
Color = c(rep(c("blue", "blue", "blue", "green"),4),
          rep(c("green", "green", "green", "blue"),4))
fisher.test(Species, Color)
cramerV(Species, Color)
```

cramerVFit

Cramer's V for chi-square goodness-of-fit tests

Description

Calculates Cramer's V for a vector of counts and expected counts.

Usage

```
cramerVFit(x, p = rep(1/length(x), length(x)), digits = 4, ...)
```

Arguments

x	A vector of observed counts.
p	A vector of expected or default probabilities.
digits	The number of significant digits in the output.
...	Additional arguments passed to <code>chisq.test</code> .

Details

In the case of single vector of counts and expected probabilities, a modification of Cramer's V can be used to indicate the degree of deviation from the expected probabilities.

It is not affected by sample size and can be used as an effect size.

In the case of equally-distributed expected frequencies, Cramer's V will be equal to 1 when all counts are in one category, and it will be equal to 0 when the counts are equally distributed across categories.

Value

A single statistic, Cramer's V.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_03.html

See Also

[CramerV](#)

Examples

```
### Equal probabilities example
### From http://rcompanion.org/handbook/H_03.html
nail.color = c("Red", "None", "White", "Green", "Purple", "Blue")
observed = c( 19, 3, 1, 1, 2, 2 )
expected = c( 1/6, 1/6, 1/6, 1/6, 1/6, 1/6 )
chisq.test(x = observed, p = expected)
cramerVFit(x = observed, p = expected)

### Unequal probabilities example
### From http://rcompanion.org/handbook/H_03.html
race = c("White", "Black", "American Indian", "Asian", "Pacific Islander",
        "Two or more races")
observed = c(20, 9, 9, 1, 1, 1)
expected = c(0.775, 0.132, 0.012, 0.054, 0.002, 0.025)
chisq.test(x = observed, p = expected)
cramerVFit(x = observed, p = expected)

### Examples of perfect and zero fits
cramerVFit(c(100, 0, 0, 0, 0))
cramerVFit(c(10, 10, 10, 10, 10))
```

epsilonSquared

Epsilon-squared

Description

Calculates epsilon-squared for a table with one ordinal variable and one nominal variable.

Usage

```
epsilonSquared(x, g = NULL, group = "row", digits = 3, ...)
```

Arguments

x	Either a two-way table or a two-way matrix. Can also be a vector of observations of an ordinal variable.
g	If x is a vector, g is the vector of observations for the grouping, nominal variable.
group	If x is a table or matrix, group indicates whether the "row" or the "column" variable is the nominal, grouping variable.
digits	The number of significant digits in the output.
...	Additional arguments passed to the <code>kruskal.test</code> function.

Details

Epsilon-squared is used as a measure of association for the Kruskal-Wallis test or for a two-way table with one ordinal and one nominal variable.

Value

A single statistic, epsilon-squared

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_11.html

See Also

[freemanTheta](#)

Examples

```
data(Breakfast)
library(coin)
chisq_test(Breakfast, scores = list("Breakfast" = c(-2, -1, 0, 1, 2)))
epsilonSquared(Breakfast)
```

```
data(PoohPiglet)
kruskal.test(Likert ~ Speaker, data = PoohPiglet)
epsilonSquared(x = PoohPiglet$Likert, g = PoohPiglet$Speaker)
```

freemanTheta

Freeman's theta

Description

Calculates Freeman's theta for a table with one ordinal variable and one nominal variable.

Usage

```
freemanTheta(x, g = NULL, group = "row", verbose = FALSE,  
progress = FALSE, digits = 3)
```

Arguments

x	Either a two-way table or a two-way matrix. Can also be a vector of observations of an ordinal variable.
g	If x is a vector, g is the vector of observations for the grouping, nominal variable.
group	If x is a table or matrix, group indicates whether the "row" or the "column" variable is the nominal, grouping variable.
verbose	If TRUE, prints statistics for each comparison.
progress	If TRUE, prints a message as each comparison is conducted.
digits	The number of significant digits in the output.

Details

Freeman's coefficient of differentiation (theta) is used as a measure of association for a two-way table with one ordinal and one nominal variable. See Freeman (1965).

Value

A single statistic, Freeman's theta

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

Freeman, L.C. 1965. Elementary Applied Statistics for Students in Behavioral Science. Wiley.
http://rcompanion.org/handbook/H_11.html

See Also

[epsilonSquared](#)

Examples

```

data(Breakfast)
library(coin)
chisq_test(Breakfast, scores = list("Breakfast" = c(-2, -1, 0, 1, 2)))
freemanTheta(Breakfast)

### Example from Freeman (1965), Table 10.6
Input =(
"Social.adjustment  5  4  3  2  1
Marital.status
Single              1  2  5  2  0
Married             10  5  5  0  0
Widowed             0  0  2  2  1
Divorced            0  0  0  2  3
")
Table = as.table(read.ftable(textConnection(Input)))
freemanTheta(Table)

data(PoohPiglet)
kruskal.test(Likert ~ Speaker, data = PoohPiglet)
freemanTheta(x = PoohPiglet$Likert, g = PoohPiglet$Speaker)

```

fullPTable

Convert a lower triangle matrix to a full matrix

Description

Converts a lower triangle matrix to a full matrix.

Usage

```
fullPTable(PT)
```

Arguments

PT A lower triangle matrix.

Details

This function is useful to convert a lower triangle matrix of p-values from a pairwise test to a full matrix. A full matrix can be passed to `multcompLetters` in the `multcompView` package to produce a compact letter display.

Value

A full matrix.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_08.html

Examples

```
### Example with pairwise.wilcox.test
data(BrendonSmall)
BrendonSmall$Instructor = factor(BrendonSmall$Instructor,
                                levels = c('Brendon Small', 'Jason Penopolis',
                                             'Paula Small', 'Melissa Robbins',
                                             'Coach McGuirk'))
P = pairwise.wilcox.test(x = BrendonSmall$Score, g = BrendonSmall$Instructor)
PT = P$p.value
PT
PT1 = fullPTable(PT)
PT1
library(multcompView)
multcompLetters(PT1)
```

groupwiseCMH

Post-hoc tests for Cochran-Mantel-Haenszel test

Description

Conducts groupwise tests of association on a three-way contingency table.

Usage

```
groupwiseCMH(x, group = 3, fisher = TRUE, gtest = FALSE, chisq = FALSE,
             method = "fdr", correct = "none", digits = 3, ...)
```

Arguments

x	A three-way contingency table.
group	The dimension of the table to use as the grouping variable. Will be 1, 2, or 3.
fisher	If TRUE, conducts Fisher exact test.
gtest	If TRUE, conducts G test of association.
chisq	If TRUE, conducts Chi-square test of association.
method	The method to use to adjust p-values. See ?p.adjust.
correct	The correction to apply to the G test. See GTest.
digits	The number of digits for numbers in the output.
...	Other arguments passed to chisq.test or GTest.

Details

If more than one of `fisher`, `gtest`, or `chisq` is set to `TRUE`, only one type of test of association will be conducted.

Value

A data frame of groups, test used, p-values, and adjusted p-values.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_06.html

See Also

[nominalSymmetryTest](#), [pairwiseMcNemar](#), [pairwiseNominalIndependence](#), [pairwiseNominalMatrix](#)

Examples

```
### Post-hoc for Cochran-Mantel-Haenszel test
data(AndersonBias)
Table = xtabs(Count ~ Sex + Result + County,
             data=AndersonBias)
ftable(Table)
mantelhaen.test(Table)
groupwiseCMH(Table,
              group = 3,
              fisher = TRUE,
              gtest = FALSE,
              chisq = FALSE,
              method = "fdr",
              correct = "none",
              digits = 3)
```

groupwiseGeometric *Groupwise geometric means and confidence intervals*

Description

Calculates geometric means and confidence intervals for groups.

Usage

```
groupwiseGeometric(formula = NULL, data = NULL, var = NULL,
                  group = NULL, conf = 0.95, na.rm = TRUE, digits = 3, ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	The measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)
conf	The confidence interval to use.
na.rm	If TRUE, removes NA values in the measurement variable.
digits	The number of significant figures to use in output.
...	Other arguments. Not currently useful.

Details

The input should include either formula and data; or data, var, and group. (See examples).

The function computes means, standard deviations, standard errors, and confidence intervals on log-transformed values. Confidence intervals are calculated in the traditional manner with the t-distribution. These statistics assume that the data are log-normally distributed. For data not meeting this assumption, medians and confidence intervals by bootstrap may be more appropriate.

Value

A data frame of geometric means, standard deviations, standard errors, and confidence intervals.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Results for ungrouped (one-sample) data can be obtained by either setting the right side of the formula to 1, e.g. $y \sim 1$, or by setting `group=NULL`.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/C_03.html

See Also

[groupwiseMean](#), [groupwiseMedian](#) [groupwiseHuber](#)

Examples

```
### Example with formula notation
data(Catbus)
groupwiseGeometric(Steps ~ Sex + Teacher,
                   data = Catbus)

### Example with variable notation
data(Catbus)
groupwiseGeometric(data = Catbus,
                   var   = "Steps",
                   group = c("Sex", "Teacher"))
```

groupwiseHuber

Groupwise Huber M-estimators and confidence intervals

Description

Calculates Huber M-estimator and confidence intervals for groups.

Usage

```
groupwiseHuber(formula = NULL, data = NULL, var = NULL, group = NULL,
               conf.level = 0.95, ci.type = "wald", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	The measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)
conf.level	The confidence interval to use.
ci.type	The type of confidence interval to use. Can be "wald" or "boot". See HuberM for details.
...	Other arguments passed to the HuberM function.

Details

A wrapper for the [HuberM](#) function to allow easy output for multiple groups.

The input should include either formula and data; or data, var, and group. (See examples).

Value

A data frame of requested statistics by group.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Results for ungrouped (one-sample) data can be obtained by either setting the right side of the formula to 1, e.g. $y \sim 1$, or by setting `group=NULL`.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/rcompanion/d_08a.html

See Also

[groupwiseMean](#), [groupwiseMedian](#) [groupwiseGeometric](#)

Examples

```
### Example with formula notation
data(Catbus)
groupwiseHuber(Steps ~ Teacher + Sex,
               data      = Catbus,
               ci.type = "wald")

### Example with variable notation
data(Catbus)
groupwiseHuber(data      = Catbus,
               var       = "Steps",
               group     = c("Teacher", "Sex"),
               ci.type = "wald")
```

groupwiseMean

Groupwise means and confidence intervals

Description

Calculates means and confidence intervals for groups.

Usage

```
groupwiseMean(formula = NULL, data = NULL, var = NULL, group = NULL,
              conf = 0.95, R = 5000, boot = FALSE, traditional = TRUE,
              normal = FALSE, basic = FALSE, percentile = FALSE, bca = FALSE,
              digits = 3, ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	The measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)
conf	The confidence interval to use.
R	The number of bootstrap replicates to use for bootstrapped statistics.
boot	If TRUE, includes the mean of the bootstrapped means. This can be used as an estimate of the mean for the group.
traditional	If TRUE, includes the traditional confidence intervals for the group means, using the t-distribution.
normal	If TRUE, includes the normal confidence intervals for the group means by bootstrap. See boot.ci .
basic	If TRUE, includes the basic confidence intervals for the group means by bootstrap. See boot.ci .
percentile	If TRUE, includes the percentile confidence intervals for the group means by bootstrap. See boot.ci .
bca	If TRUE, includes the BCa confidence intervals for the group means by bootstrap. See boot.ci .
digits	The number of significant figures to use in output.
...	Other arguments passed to the boot function.

Details

The input should include either `formula` and `data`; or `data`, `var`, and `group`. (See examples).

With some options, the function may not handle missing values well. This seems to happen particularly with `bca = TRUE`.

Value

A data frame of requested statistics by group.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Results for ungrouped (one-sample) data can be obtained by either setting the right side of the formula to 1, e.g. $y \sim 1$, or by setting `group=NULL`.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/C_03.html

See Also

[groupwiseMedian](#), [groupwiseHuber](#) [groupwiseGeometric](#)

Examples

```
### Example with formula notation
data(Catbus)
groupwiseMean(Steps ~ Teacher + Sex,
              data      = Catbus,
              traditional = FALSE,
              percentile = TRUE)

### Example with variable notation
data(Catbus)
groupwiseMean(data      = Catbus,
              var        = "Steps",
              group      = c("Teacher", "Sex"),
              traditional = FALSE,
              percentile = TRUE)
```

groupwiseMedian	<i>Groupwise medians and confidence intervals</i>
-----------------	---

Description

Calculates medians and confidence intervals for groups.

Usage

```
groupwiseMedian(formula = NULL, data = NULL, var = NULL, group = NULL,
                conf = 0.95, R = 5000, boot = FALSE, pseudo = FALSE, basic = FALSE,
                normal = FALSE, percentile = FALSE, bca = TRUE, wilcox = FALSE,
                exact = FALSE, digits = 3, ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	The measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)

conf	The confidence interval to use.
R	The number of bootstrap replicates to use for bootstrapped statistics.
boot	If TRUE, includes the mean of the bootstrapped medians. This can be used as an estimate of the median for the group.
pseudo	If TRUE, includes the pseudo median from <code>wilcox.test</code> .
basic	If TRUE, includes the basic confidence intervals for the group means by bootstrap. See boot.ci .
normal	If TRUE, includes the normal confidence intervals for the group means by bootstrap. See boot.ci .
percentile	If TRUE, includes the percentile confidence intervals for the group means by bootstrap. See boot.ci .
bca	If TRUE, includes the BCa confidence intervals for the group means by bootstrap. See boot.ci .
wilcox	If TRUE, includes the wilcox confidence intervals from <code>wilcox.test</code> .
exact	If TRUE, includes the "exact" confidence intervals from <code>MedianCI</code> .
digits	The number of significant figures to use in output.
...	Other arguments passed to the <code>boot</code> function.

Details

The input should include either formula and data; or data, var, and group. (See examples).

With some options, the function may not handle missing values well. This seems to happen particularly with `bca = TRUE`.

Value

A data frame of requested statistics by group.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Results for ungrouped (one-sample) data can be obtained by either setting the right side of the formula to 1, e.g. `y ~ 1`, or by setting `group=NULL`.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/E_04.html

See Also

[groupwiseMean](#), [groupwiseHuber](#) [groupwiseGeometric](#)

Examples

```

### Example with formula notation
data(Catbus)
groupwiseMedian(Steps ~ Teacher + Sex,
                 data      = Catbus,
                 bca       = FALSE,
                 percentile = TRUE,
                 R         = 1000)

### Example with variable notation
data(Catbus)
groupwiseMedian(data      = Catbus,
                 var       = "Steps",
                 group     = c("Teacher", "Sex"),
                 bca       = FALSE,
                 percentile = TRUE,
                 R         = 1000)

```

groupwisePercentile *Groupwise percentiles and confidence intervals*

Description

Calculates percentiles and confidence intervals for groups.

Usage

```

groupwisePercentile(formula = NULL, data = NULL, var = NULL,
                    group = NULL, conf = 0.95, tau = 0.5, type = 7, R = 5000,
                    boot = FALSE, basic = FALSE, normal = FALSE, percentile = FALSE,
                    bca = TRUE, digits = 3, ...)

```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	If no formula is given, the measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)
conf	The confidence interval to use.
tau	The percentile to use, expressed as a quantile, e.g. 0.5 for median, 0.25 for 25th percentile.
type	The type value passed to the quantile function

R	The number of bootstrap replicates to use for bootstrapped statistics.
boot	If TRUE, includes the mean of the bootstrapped percentile. This can be used as an estimate of the percentile for the group.
basic	If TRUE, includes the basic confidence intervals for the group means by bootstrap. See boot.ci .
normal	If TRUE, includes the normal confidence intervals for the group means by bootstrap. See boot.ci .
percentile	If TRUE, includes the percentile confidence intervals for the group means by bootstrap. See boot.ci .
bca	If TRUE, includes the BCa confidence intervals for the group means by bootstrap. See boot.ci .
digits	The number of significant figures to use in output.
...	Other arguments passed to the boot function.

Details

The input should include either formula and data; or data, var, and group. (See examples).

With some options, the function may not handle missing values well. This seems to happen particularly with `bca = TRUE`.

Value

A data frame of requested statistics by group

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Results for ungrouped (one-sample) data can be obtained by either setting the right side of the formula to 1, e.g. `y ~ 1`, or by setting `group=NULL`.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_15.html

See Also

[groupwiseMean](#), [groupwiseHuber](#), [groupwiseGeometric](#), [groupwiseMedian](#)

Examples

```

### Example with formula notation
data(Catbus)
groupwisePercentile(Steps ~ Teacher + Sex,
                    data      = Catbus,
                    tau       = 0.25,
                    bca       = FALSE,
                    percentile = TRUE,
                    R         = 1000)

### Example with variable notation
data(Catbus)
groupwisePercentile(data      = Catbus,
                    var       = "Steps",
                    group     = c("Teacher", "Sex"),
                    tau       = 0.25,
                    bca       = FALSE,
                    percentile = TRUE,
                    R         = 1000)

```

groupwiseSum

Groupwise sums

Description

Calculates sums for groups.

Usage

```
groupwiseSum(formula = NULL, data = NULL, var = NULL, group = NULL,
             digits = 3, ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
var	The measurement variable to use. The name is in double quotes.
group	The grouping variable to use. The name is in double quotes. Multiple names are listed as a vector. (See example.)
digits	The number of significant figures to use in output.
...	Other arguments passed to the sum function

Details

The input should include either formula and data; or data, var, and group. (See examples).

Value

A data frame of statistics by group.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The variables on the right side are used for the grouping variables.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

See Also

[groupwiseMean](#), [groupwiseMedian](#), [groupwiseHuber](#), [groupwiseGeometric](#)

Examples

```
### Example with formula notation
data(AndersonBias)
groupwiseSum(Count ~ Result + Sex,
             data      = AndersonBias)

### Example with variable notation
data(AndersonBias)
groupwiseSum(data      = AndersonBias,
             var        = "Count",
             group      = c("Result", "Sex"))
```

HayleySmith

Hypothetical data for responses about adopting lawn care practices

Description

A data frame in long form with yes/no responses for four lawn care practices for each of 14 respondents. Hypothetical data.

Usage

```
HayleySmith
```

Format

An object of class `data.frame` with 56 rows and 3 columns.

Source

http://rcompanion.org/handbook/H_05.html

 Monarchs

Hypothetical data for monarch butterflies in gardens

Description

A data frame of the number of monarch butterflies in three gardens. Hypothetical data.

Usage

```
Monarchs
```

Format

An object of class `data.frame` with 24 rows and 2 columns.

Source

http://rcompanion.org/handbook/J_01.html

 nagelkerke

Pseudo r-squared measures for various models

Description

Produces McFadden, Cox and Snell, and Nagelkerke pseudo R-squared measures, along with p-values, for models.

Usage

```
nagelkerke(fit, null = NULL, restrictNobs = FALSE)
```

Arguments

<code>fit</code>	The fitted model object for which to determine pseudo r-squared.
<code>null</code>	The null model object against which to compare the fitted model object. The null model must be nested in the fitted model to be valid. Specifying the null is optional for some model object types and is required for others.
<code>restrictNobs</code>	If TRUE, limits the observations for the null model to those used in the fitted model. Works with only some model object types.

Details

Pseudo R-squared values are not directly comparable to the R-squared for OLS models. Nor can they be interpreted as the proportion of the variability in the dependent variable that is explained by model. Instead pseudo R-squared measures are relative measures among similar models indicating how well the model explains the data.

Cox and Snell is also referred to as ML. Nagelkerke is also referred to as Cragg and Uhler.

Model objects accepted are `lm`, `glm`, `gls`, `lme`, `lmer`, `lmerTest`, `nls`, `clm`, `clmm`, `vglm`, `glmer`, `negbin`, `zeroinfl`, `betareg`, and `rq`.

Model objects that require the null model to be defined are `nls`, `lmer`, `glmer`, and `clmm`. Other objects use the `update` function to define the null model.

Likelihoods are found using ML (REML = FALSE).

The fitted model and the null model should be properly nested. That is, the terms of one need to be a subset of the the other, and they should have the same set of observations. One issue arises when there are NA values in one variable but not another, and observations with NA are removed in the model fitting. The result may be fitted and null models with different sets of observations. Setting `restrictNobs` to TRUE ensures that only observations in the fit model are used in the null model. This appears to work for `lm` and some `glm` models, but causes the function to fail for other model object types.

Some pseudo R-squared measures may not be appropriate or useful for some model types.

Calculations are based on log likelihood values for models. Results may be different than those based on deviance.

Value

A list of six objects describing the models used, the pseudo r-squared values, the likelihood ratio test for the model, the number of obervaton for the models, messages, and any warnings.

Acknowledgements

My thanks to Jan-Herman Kuiper of Keele University for suggesting the `restrictNobs` fix.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/G_10.html

See Also

[accuracy nagelkerkeHermite](#)

Examples

```

### Logistic regression example
data(AndersonBias)
model = glm(Result ~ County + Sex + County:Sex,
            weight = Count,
            data = AndersonBias,
            family = binomial(link="logit"))
nagelkerke(model)

### Quadratic plateau example
### With nls, the null needs to be defined
data(BrendonSmall)
quadplat = function(x, a, b, clx) {
  ifelse(x < clx, a + b * x + (-0.5*b/clx) * x * x,
        a + b * clx + (-0.5*b/clx) * clx * clx)}
model = nls(Sodium ~ quadplat(Calories, a, b, clx),
            data = BrendonSmall,
            start = list(a = 519,
                        b = 0.359,
                        clx = 2304))
nullfunct = function(x, m){m}
null.model = nls(Sodium ~ nullfunct(Calories, m),
                data = BrendonSmall,
                start = list(m = 1346))
nagelkerke(model, null=null.model)

```

nagelkerkeHermite

Pseudo r-squared measures for hermite models

Description

Produces McFadden, Cox and Snell, and Nagelkerke pseudo R-squared measures, along with p-value for the model, for hermite regression objects.

Usage

```
nagelkerkeHermite(fit, null)
```

Arguments

fit	The fitted model object for which to determine pseudo r-squared.
null	The null model object against which to compare the fitted model object. The null model must be nested in the fitted model to be valid.

Details

Hermite regression is performed with the `hermite` package.

For pseudo r-squared measures, Cox and Snell is also referred to as ML. Nagelkerke is also referred to as Cragg and Uhler.

The fit model and the null model should be properly nested. That is, the terms of one need to be a subset of the the other, and they should have the same set of observations.

Value

A list of six objects describing the models used, the pseudo r-squared values, the likelihood ratio test for the model, the AIC for the fitted and null models, the number of observations for the models, and any warnings.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/J_01.html

See Also

[nagelkerke](#)

Examples

```
data(Monarchs)
library(hermite)
model = glm.hermite(Monarchs ~ Garden,
                   data = Monarchs,
                   link = "log",
                   m=3)
null = glm.hermite(Monarchs ~ 1,
                  data = Monarchs,
                  link = "log",
                  m=3)
nagelkerkeHermite(model, null)
```

nominalSymmetryTest *Exact and Monte Carlo symmetry tests for paired contingency tables*

Description

Conducts an omnibus symmetry test for a paired contingency table and then post-hoc pairwise tests. This is similar to McNemar and McNemar-Bowker tests in use.

Usage

```
nominalSymmetryTest(x, method = "fdr", digits = 3, ...)
```

Arguments

x	A two-way contingency table. It must be square. It can have two or more levels for each dimension.
method	The method to adjust multiple p-values. See p.adjust .
digits	The number of significant digits in the output.
...	Additional arguments, passed to multinomial.test .

Details

If Monte Carlo is not used, the test of symmetry uses an exact test by conducting either a binomial or multinomial goodness-of-fit test.

These are equivalent to uncorrected McNemar and McNemar-Bowker tests, but will not fail when there are zeros in critical cells, as will the `mcnemar.test` function.

Value

A list containing: a data frame of results of the global test; a data frame of results of the pairwise results; and a data frame mentioning the p-value adjustment method.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_05.html

See Also

[pairwiseMcnemar](#), [groupwiseCMH](#), [pairwiseNominalIndependence](#), [pairwiseNominalMatrix](#)

Examples

```
### 2 x 2 repeated matrix example
data(AndersonRainBarrel)
nominalSymmetryTest(AndersonRainBarrel)

### 3 x 3 repeated matrix example with Monte Carlo
data(AndersonRainGarden)
nominalSymmetryTest(AndersonRainGarden,
                    MonteCarlo = TRUE,
                    ntrial      = 10000)

### 4 x 4 repeated matrix example that fails with mcnemar.test
data(Religion)
```



```
nominalSymmetryTest(Religion,
                    MonteCarlo = TRUE,
                    ntrial      = 10000)
```

 Nurseries

Data for proportion of good practices followed by plant nurseries

Description

A data frame with two variables: size of plant nursery in hectares, and proportion of good practices followed by the nursery

Usage

```
Nurseries
```

Format

An object of class `data.frame` with 38 rows and 2 columns.

Source

Mangiafico, S.S., Newman, J.P., Mochizuki, M.J., and Zurawski, D. (2008). Adoption of sustainable practices to protect and conserve water resources in container nurseries with greenhouse facilities. *Acta horticulturae* 797, 367-372.

 pairwiseDifferences

Pairwise differences for unreplicated CBD

Description

Calculates the differences in the response variable for each pair of levels of a grouping variable in an unreplicated complete block design.

Usage

```
pairwiseDifferences(formula = NULL, data = NULL, x = NULL, g = NULL,
                   plotit = FALSE, factorize = FALSE)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. $y \sim \text{group}$.
data	The data frame to use.
x	The vector of the response variable.
g	The vector of the grouping variable.
plotit	If TRUE, then produces bar plots of the differences.
factorize	If TRUE, then adds a column to the output data frame consisting of the differences as a factor variable. This output is added automatically if <code>plotit = TRUE</code> .

Details

The main use of the function is to check the shape of the distribution of differences in responses for paired t-test, paired rank-sum test, Friedman test, or Quade test.

The function assumes that the data frame is already ordered by the blocking variable, so that the first observation of Group 1 is paired with the first observation of Group 2, and so on.

The function assumes that the data are in complete block design. That is, for any level of the grouping variable in Group 1 there exists one paired value in Group 2, and so on.

The input should include either formula and data; or x, and g.

Value

A data frame of the paired groups, the differences in their response variables, and optionally the differences expressed as a factor variable. If `plotit = TRUE`, then also produce one or more plots.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_10.html

Examples

```
### Two-sample paired data example with bar plot
data(Pooh)
Pooh.diff = pairwiseDifferences(Likert ~ Time,
                               data=Pooh,
                               plotit = TRUE)

### Unreplicated complete block design example with bar plots
```

```

data(BobBelcher)
Bob.diff = pairwiseDifferences(Likert ~ Instructor,
                              data=BobBelcher,
                              factorize=TRUE)

library(lattice)
histogram(~ Difference.f | Comparison,
          data=Bob.diff,
          type = "count",
          layout=c(2,5))

```

pairwiseMcnemar

Pairwise McNemar and related tests for Cochran Q test post-hoc

Description

Conducts pairwise McNemar, exact, and permutation tests as a post-hoc to Cochran Q test.

Usage

```

pairwiseMcnemar(formula = NULL, data = NULL, x = NULL, g = NULL,
                block = NULL, test = "exact", method = "fdr", digits = 3,
                correct = FALSE)

```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. $y \sim \text{group} \mid \text{block}$.
data	The data frame to use.
x	The response variable.
g	The grouping variable.
block	The blocking variable.
test	If "exact", conducts an exact test of symmetry analogous to a McNemar test. If "mcnemar", conducts a McNemar test of symmetry. If "permutation", conducts a permutation test analogous to a McNemar test.
method	The method for adjusting multiple p-values. See p.adjust .
digits	The number of significant digits in the output.
correct	If TRUE, applies a continuity correction for the McNemar test.

Details

The component tables for the pairwise tests must be of size 2×2 .

The input should include either formula and data; or x, g, and block.

Value

A list containing: a data frame of results of the global test; a data frame of results of the pairwise results; and a data frame mentioning the p-value adjustment method.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable. The second variable on the right side is used for the blocking variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_07.html

See Also

[nominalSymmetryTest](#), [groupwiseCMH](#), [pairwiseNominalIndependence](#), [pairwiseNominalMatrix](#)

Examples

```
### Cochran Q post-hoc example
data(HayleySmith)
library(RVAideMemoire)
cochran.qtest(Response ~ Practice | Student,
              data = HayleySmith)
HayleySmith$Practice = factor(HayleySmith$Practice,
                              levels = c("MowHeight", "SoilTest",
                                          "Clippings", "Irrigation"))
PT = pairwiseMcnemar(Response ~ Practice | Student,
                    data = HayleySmith,
                    test = "exact",
                    method = "fdr",
                    digits = 3)
PT
PT = PT$Pairwise
cldList(comparison = PT$Comparison,
        p.value = PT$p.adjust,
        threshold = 0.05)
```

pairwiseMedianMatrix *Pairwise Mood's median tests with matrix output*

Description

Conducts pairwise Mood's median tests across groups.

Usage

```
pairwiseMedianMatrix(formula = NULL, data = NULL, x = NULL, g = NULL,
  exact = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
exact	If TRUE, then asks the <code>mood.medtest</code> function to conduct an exact test. If NULL, then uses an exact test if the number of values is less than 200. See mood.medtest .
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to mood.medtest .

Details

The input should include either `formula` and `data`; or `x`, and `g`.

Mood's median test compares medians among two or more groups. See http://rcompanion.org/handbook/F_09.html for further discussion of this test.

The `pairwiseMedianMatrix` function can be used as a post-hoc method following an omnibus Mood's median test. The matrix output can be converted to a compact letter display.

Value

A list consisting of: a matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_09.html

See Also

[pairwiseMedianTest](#)

Examples

```
data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                           levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseMedianMatrix(Likert ~ Speaker,
                          data = PoohPiglet,
                          exact = NULL,
                          method = "fdr")$Adjusted

PT
library(multcompView)
multcompLetters(PT,
                compare="<",
                threshold=0.05,
                Letters=letters)
```

pairwiseMedianTest *Pairwise Mood's median tests*

Description

Conducts pairwise Mood's median tests across groups.

Usage

```
pairwiseMedianTest(formula = NULL, data = NULL, x = NULL, g = NULL,
                   exact = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
exact	If TRUE, then asks the <code>mood.medtest</code> function to conduct an exact test. If NULL, then uses an exact test if the number of values is less than 200. See mood.medtest .
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to <code>codemood.medtest</code> .

Details

The input should include either formula and data; or x, and g.

Mood's median test compares medians among two or more groups. See http://rcompanion.org/handbook/F_09.html for further discussion of this test.

The pairwiseMedianTest function can be used as a post-hoc method following an omnibus Mood's median test.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_09.html

See Also

[pairwiseMedianMatrix](#)

Examples

```
data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                           levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseMedianTest(Likert ~ Speaker,
                        data = PoohPiglet,
                        exact = NULL,
                        method = "fdr")

PT
cldList(comparison = PT$Comparison,
        p.value = PT$p.adjust,
        threshold = 0.05)
```

pairwiseModelAnova *Compare model objects with F test and likelihood ratio test*

Description

Compares a series of models with pairwise F tests and likelihood ratio tests.

Usage

```
pairwiseModelAnova(fits, ...)
```

Arguments

fits	A series of model object names, separated by commas.
...	Other arguments passed to list.

Details

For comparisons to be valid, both models must have the same data, without transformations, use the same dependent variable, and be fit with the same method.

To be valid, models need to be nested.

Value

A list of: The calls of the models compared; a data frame of comparisons and F tests; and a data frame of comparisons and likelihood ratio tests.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

See Also

[compareGLM](#), [compareLM](#)

Examples

```
### Compare among polynomial models
data(BrendonSmall)
BrendonSmall$Calories = as.numeric(BrendonSmall$Calories)

BrendonSmall$Calories2 = BrendonSmall$Calories * BrendonSmall$Calories
BrendonSmall$Calories3 = BrendonSmall$Calories * BrendonSmall$Calories *
  BrendonSmall$Calories
BrendonSmall$Calories4 = BrendonSmall$Calories * BrendonSmall$Calories *
  BrendonSmall$Calories * BrendonSmall$Calories
model.1 = lm(Sodium ~ Calories, data = BrendonSmall)
model.2 = lm(Sodium ~ Calories + Calories2, data = BrendonSmall)
```



```

model.3 = lm(Sodium ~ Calories + Calories2 + Calories3, data = BrendonSmall)
model.4 = lm(Sodium ~ Calories + Calories2 + Calories3 + Calories4,
             data = BrendonSmall)
pairwiseModelAnova(model.1, model.2, model.3, model.4)

```

pairwiseNominalIndependence

Pairwise tests of independence for nominal data

Description

Conducts pairwise tests for a 2-dimensional matrix, in which at least one dimension has more than two levels, as a post-hoc test. Conducts Fisher exact, Chi-square, or G-test.

Usage

```

pairwiseNominalIndependence(x, compare = "row", fisher = TRUE,
                             gtest = TRUE, chisq = TRUE, method = "fdr", correct = "none",
                             digits = 3, ...)

```

Arguments

x	A two-way contingency table. At least one dimension should have more than two levels.
compare	If "row", treats the rows as the grouping variable. If "column", treats the columns as the grouping variable.
fisher	If "TRUE", conducts fisher exact test.
gtest	If "TRUE", conducts G-test.
chisq	If "TRUE", conducts Chi-square test of association.
method	The method to adjust multiple p-values. See p.adjust .
correct	The correction method to pass to GTest .
digits	The number of significant digits in the output.
...	Additional arguments, passed to fisher.test , GTest , or chisq.test .

Value

A data frame of comparisons, p-values, and adjusted p-values.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_04.html

See Also

[pairwiseMcNemar](#), [groupwiseCMH](#), [nominalSymmetryTest](#), [pairwiseNominalMatrix](#)

Examples

```
### Independence test for a 4 x 2 matrix
data(Anderson)
fisher.test(Anderson)
Anderson = Anderson[(c("Heimlich", "Bloom", "Dougal", "Cobblestone"),)]
PT = pairwiseNominalIndependence(Anderson,
                                fisher = TRUE,
                                gtest = FALSE,
                                chisq = FALSE)

PT
cldList(comparison = PT$Comparison,
        p.value     = PT$p.adj.Fisher,
        threshold  = 0.05)
```

pairwiseNominalMatrix *Pairwise tests of independence for nominal data with matrix output*

Description

Conducts pairwise tests for a 2-dimensional matrix, in which at least one dimension has more than two levels, as a post-hoc test. Conducts Fisher exact, Chi-square, or G-test.

Usage

```
pairwiseNominalMatrix(x, compare = "row", fisher = TRUE, gtest = FALSE,
                      chisq = FALSE, method = "fdr", correct = "none", digits = 3, ...)
```

Arguments

x	A two-way contingency table. At least one dimension should have more than two levels.
compare	If "row", treats the rows as the grouping variable. If "column", treats the columns as the grouping variable.
fisher	If "TRUE", conducts fisher exact test.
gtest	If "TRUE", conducts G-test.
chisq	If "TRUE", conducts Chi-square test of association.
method	The method to adjust multiple p-values. See p.adjust .
correct	The correction method to pass to GTest .
digits	The number of significant digits in the output.
...	Additional arguments, passed to fisher.test , GTest , or chisq.test .

Arguments

x	A two-way contingency table. One dimension is ordered and one is non-ordered nominal.
compare	If "row", treats the rows as the grouping variable. If "column", treats the columns as the grouping variable.
scores	Optional vector to specify the spacing of the ordered variable.
method	The method to adjust multiple p-values. See p.adjust .
digits	The number of significant digits in the output.
...	Additional arguments, passed to chisq_test .

Value

A data frame of comparisons, p-values, and adjusted p-values.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/H_09.html

See Also

[pairwiseNominalIndependence](#)

Examples

```
### Independence test for table with one ordered variable
data(Breakfast)
require(coin)
chisq_test(Breakfast,
           scores = list("Breakfast" = c(-2, -1, 0, 1, 2)))
PT = pairwiseOrdinalIndependence(Breakfast, compare = "row")
PT
cldList(comparison = PT$Comparison,
        p.value     = PT$p.value,
        threshold   = 0.05)

### Similar to Kruskal-Wallis test for Likert data
data(PoohPiglet)
XT = xtabs(~ Speaker + Likert, data = PoohPiglet)
XT
require(coin)
chisq_test(XT,
           scores = list("Likert" = c(1, 2, 3, 4, 5)))
PT=pairwiseOrdinalIndependence(XT, compare = "row")
PT
cldList(comparison = PT$Comparison,
```

```
p.value = PT$p.value,  
threshold = 0.05)
```

pairwiseOrdinalMatrix *Pairwise two-sample ordinal regression with matrix output*

Description

Performs pairwise two-sample ordinal regression across groups.

Usage

```
pairwiseOrdinalMatrix(formula = NULL, data = NULL, x = NULL, g = NULL,  
  method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to clm .

Details

Ordinal regression is analogous to general linear regression or generalized linear regression for cases where the dependent variable is an ordinal variable. The `ordinal` package provides a flexible and powerful implementation of ordinal regression.

The `pairwiseOrdinalMatrix` function can be used as a post-hoc method following an omnibus ordinal regression whose form is analogous to a one-way analysis of variance. The matrix output can be converted to a compact letter display.

The `x` variable must be an ordered factor.

Value

A list consisting of: A matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/G_07.html

See Also

[pairwiseOrdinalTest](#)

Examples

```
data(PoohPiglet)
PoohPiglet$Likert.f = factor(PoohPiglet$Likert, ordered = TRUE)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                             levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseOrdinalMatrix(Likert.f ~ Speaker,
                             data = PoohPiglet,
                             method = "fdr")$Adjusted

PT
library(multcompView)
multcompLetters(PT,
                 compare="<",
                 threshold=0.05,
                 Letters=letters)
```

pairwiseOrdinalPairedMatrix

Pairwise two-sample ordinal regression for blocked data with matrix output

Description

Performs pairwise two-sample ordinal regression across groups for paired or blocked data.

Usage

```
pairwiseOrdinalPairedMatrix(formula = NULL, data = NULL, x = NULL,
                             g = NULL, b = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group block</code> .
data	The data frame to use.
x	The response variable as a vector.


```

PT = pairwiseOrdinalPairedMatrix(Likert.f ~ Instructor | Rater,
                                data      = BobBelcher,
                                threshold = "equidistant",
                                method    = "fdr")$Adjusted

PT
library(multcompView)
multcompLetters(PT,
                compare="<",
                threshold=0.05,
                Letters=letters)

```

pairwiseOrdinalPairedTest

Pairwise two-sample ordinal regression for paired or blocked data

Description

Performs pairwise two-sample ordinal regression across groups for paired or blocked data.

Usage

```
pairwiseOrdinalPairedTest(formula = NULL, data = NULL, x = NULL,
                          g = NULL, b = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group block</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
b	The blocking variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to clmm .

Details

The input should include either `formula` and `data`; or `x`, `g`, and `b`.

Ordinal regression is analogous to general linear regression or generalized linear regression for cases where the dependent variable is an ordinal variable. The `ordinal` package provides a flexible and powerful implementation of ordinal regression.

The `pairwiseOrdinalPairedTest` function can be used as a post-hoc method following an omnibus ordinal regression whose form is analogous to a one-way analysis of variance with random blocks.

The blocking variable is treated as a random variable.

The `x` variable must be an ordered factor.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable. The second variable on the right side is used for the blocking variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/G_08.html

See Also

[pairwiseOrdinalPairedMatrix](#)

Examples

```
data(BobBelcher)
BobBelcher$Likert.f = factor(BobBelcher$Likert, ordered = TRUE)
BobBelcher$Instructor = factor( BobBelcher$Instructor,
                               levels = c("Linda Belcher", "Louise Belcher",
                                           "Tina Belcher", "Bob Belcher",
                                           "Gene Belcher"))
PT = pairwiseOrdinalPairedTest(Likert.f ~ Instructor | Rater,
                              data      = BobBelcher,
                              threshold = "equidistant",
                              method   = "fdr")

PT
cldList(comparison = PT$Comparison,
        p.value    = PT$p.adjust,
        threshold  = 0.05)
```

pairwiseOrdinalTest *Pairwise two-sample ordinal regression*

Description

Performs pairwise two-sample ordinal regression across groups.

Usage

```
pairwiseOrdinalTest(formula = NULL, data = NULL, x = NULL, g = NULL,  
  method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to clm .

Details

The input should include either `formula` and `data`; or `x`, and `g`.

Ordinal regression is analogous to general linear regression or generalized linear regression for cases where the dependent variable is an ordinal variable. The `ordinal` package provides a flexible and powerful implementation of ordinal regression.

The `pairwiseOrdinalTest` function can be used as a post-hoc method following an omnibus ordinal regression whose form is analogous to a one-way analysis of variance.

The `x` variable must be an ordered factor.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/G_07.html

See Also

[pairwiseOrdinalMatrix](#)

Examples

```

data(PoohPiglet)
PoohPiglet$Likert.f = factor(PoohPiglet$Likert, ordered = TRUE)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                             levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseOrdinalTest(Likert.f ~ Speaker,
                          data = PoohPiglet,
                          method = "fdr")

PT
cldList(comparison = PT$Comparison,
         p.value = PT$p.adjust,
         threshold = 0.05)

```

pairwisePercentileTest

Pairwise permutation tests for percentiles

Description

Conducts pairwise permutation tests across groups for percentiles, medians, and proportion below a threshold value.

Usage

```

pairwisePercentileTest(formula = NULL, data = NULL, x = NULL, y = NULL,
                        test = "median", tau = 0.5, type = 7, threshold = NA,
                        comparison = "<", r = 1000, digits = 4, progress = "TRUE",
                        method = "fdr")

```

Arguments

formula	A formula indicating the response variable and the independent variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	If no formula is given, the response variable for one group.
y	The response variable for the other group.
test	The statistic to compare between groups. Can be "median", "percentile", "iqr", "proportion", "mean", or "variance".
tau	If "percentile" is chosen as the test, tau indicates the percentile to test. Expressed as a quantile. That is, 0.5 indicates a test for medians. 0.75 indicates a test for 75th percentiles.
type	The type value passed to the quantile function.

threshold	If "proportion" is chosen as the test, threshold indicates the value of the dependent variable to use as the threshold. For example, to test if there is a different in the proportion of observations below \$10,000, threshold = 10000 would be used.
comparison	If "proportion" is chosen as the test, comparison indicates the inequality to use. Options are "<", "<=", ">", ">=", or, "=="
r	The number of replicates in the permutation test.
digits	The number of significant digits in the output.
progress	If TRUE, prints a dot for every 1 percent of the progress while conducting the test.
method	The p-value adjustment method to use for multiple tests. See p.adjust .

Details

The function conducts pairwise tests using the `percentileTest` function. The user can consult the documentation for that function for additional details.

The input should include either formula and data; or x, and y.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_15.html

See Also

[percentileTest](#), [groupwisePercentile](#)

Examples

```
## Not run:
data(BrendonSmall)
PT = pairwisePercentileTest(Sodium ~ Instructor,
                           data = BrendonSmall,
                           test = "percentile",
                           tau = 0.75)

PT
cldList(p.adjust ~ Comparison,
```

```

      data      = PT,
      threshold = 0.05)

data(BrendonSmall)
PT = pairwisePercentileTest(Sodium ~ Instructor,
                           data      = BrendonSmall,
                           test      = "proportion",
                           threshold = 1300)

PT
cldList(p.adjust ~ Comparison,
        data      = PT,
        threshold = 0.05)

## End(Not run)

```

pairwisePermutationMatrix

Pairwise two-sample permutation tests with matrix output

Description

Conducts pairwise two-sample permutation tests across groups.

Usage

```
pairwisePermutationMatrix(formula = NULL, data = NULL, x = NULL,
                          g = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See <code>p.adjust</code> .
...	Additional arguments passed to <code>independence_test</code> .

Details

The input should include either formula and data; or x, and g.

Permutation tests are non-parametric tests that do not assume normally-distributed errors. See http://rcompanion.org/rcompanion/d_06a.html for further discussion of this test.

The `pairwisePermutationTest` function can be used as a post-hoc method following an omnibus permutation test analogous to a one-way analysis of variance. The matrix output can be converted to a compact letter display.

Value

A list consisting of: A matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/K_02.html

See Also

[pairwisePermutationTest](#)

Examples

```
data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                           levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwisePermutationMatrix(Likert ~ Speaker,
                              data = PoohPiglet,
                              method = "fdr")

PT
PT = PT$Adjusted
library(multcompView)
multcompLetters(PT,
                compare="<",
                threshold=0.05,
                Letters=letters)
```

pairwisePermutationSymmetry

Pairwise two-sample permutation symmetry tests

Description

Conducts pairwise two-sample permutation tests of symmetry across groups.

Usage

```
pairwisePermutationSymmetry(formula = NULL, data = NULL, x = NULL,  
  g = NULL, b = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group block</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
b	The blocking variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to symmetry_test .

Details

The input should include either formula and data; or x, g, and b.

Permutation tests are non-parametric tests that do not assume normally-distributed errors. See http://rcompanion.org/rcompanion/d_06a.html for further discussion of this test.

The `pairwisePermutationSymmetry` function can be used as a post-hoc method following an omnibus permutation test analogous to a paired one-way analysis of variance.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable. The second variable on the right side is used for the blocking variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/K_03.html

See Also

[pairwisePermutationSymmetryMatrix](#)

Examples

```

data(BobBelcher)
BobBelcher$Instructor = factor( BobBelcher$Instructor,
                               levels = c("Linda Belcher", "Louise Belcher",
                                           "Tina Belcher", "Bob Belcher",
                                           "Gene Belcher"))
BobBelcher$Likert.f = factor(BobBelcher$Likert, ordered=TRUE)
PT = pairwisePermutationSymmetry(Likert.f ~ Instructor | Rater,
                                data      = BobBelcher,
                                method   = "fdr")

PT
cldList(comparison = PT$Comparison,
        p.value    = PT$p.adjust,
        threshold  = 0.05)

```

```
pairwisePermutationSymmetryMatrix
```

Pairwise two-sample permutation symmetry tests with matrix output

Description

Conducts pairwise two-sample permutation tests for symmetry across groups.

Usage

```
pairwisePermutationSymmetryMatrix(formula = NULL, data = NULL, x = NULL,
                                  g = NULL, b = NULL, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
b	The blocking variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to symmetry_test .

Details

The input should include either formula and data; or x, g, and b.

Permutation tests are non-parametric tests that do not assume normally-distributed errors. See http://rcompanion.org/rcompanion/d_06a.html for further discussion of this test.

The `pairwisePermutationSymmetryMatrix` function can be used as a post-hoc method following an omnibus permutation test analogous to a paired one-way analysis of variance. The matrix output can be converted to a compact letter display.

Value

A list consisting of: A matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable. The second variable on the right side is used for the blocking variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/K_03.html

See Also

[pairwisePermutationSymmetry](#)

Examples

```
data(BobBelcher)
BobBelcher$Instructor = factor( BobBelcher$Instructor,
                                levels = c("Linda Belcher", "Louise Belcher",
                                             "Tina Belcher", "Bob Belcher",
                                             "Gene Belcher"))
BobBelcher$Likert.f = factor(BobBelcher$Likert, ordered=TRUE)
PT = pairwisePermutationSymmetryMatrix(Likert.f ~ Instructor | Rater,
                                       data = BobBelcher,
                                       method = "fdr")$Adjusted

PT
library(multcompView)
multcompLetters(PT,
                compare="<",
                threshold=0.05,
                Letters=letters)
```

pairwisePermutationTest

Pairwise two-sample permutation tests

Description

Conducts pairwise two-sample permutation tests across groups.

Usage

```
pairwisePermutationTest(formula = NULL, data = NULL, x = NULL, g = NULL,  
  method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to independence_test .

Details

The input should include either formula and data; or x, and g.

Permutation tests are non-parametric tests that do not assume normally-distributed errors. See http://rcompanion.org/rcompanion/d_06a.html for further discussion of this test.

The `pairwisePermutationTest` function can be used as a post-hoc method following an omnibus permutation test analogous to a one-way analysis of variance.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/K_02.html

See Also

[pairwisePermutationMatrix](#)

Examples

```

data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
  levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwisePermutationTest(Likert ~ Speaker,
  data = PoohPiglet,
  method = "fdr")

PT
cldList(comparison = PT$Comparison,
  p.value = PT$p.adjust,
  threshold = 0.05)

```

pairwiseRobustMatrix *Pairwise two-sample robust tests with matrix output*

Description

Performs pairwise two-sample robust tests across groups with matrix output.

Usage

```

pairwiseRobustMatrix(formula = NULL, data = NULL, x = NULL, g = NULL,
  est = "mom", nboot = 599, method = "fdr", ...)

```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
est	Estimate used for group comparisons. "onestep", "mom", "median", or "mean". See pb2gen for details.
nboot	The number of bootstrap samples.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to pb2gen .

Details

The input should include either formula and data; or x, and g.

The WRS2 package provides functions for robust estimation and hypothesis testing. This function invokes the pb2gen to make pairwise comparisons among groups.

The pairwiseRobustMatrix function can be used as a post-hoc method following an omnibus one-way anova with robust estimation.

Value

A list consisting of: A matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/rcompanion/d_08a.html

See Also

[pairwiseRobustTest](#)

Examples

```
data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                           levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseRobustMatrix(Likert ~ Speaker,
                          data = PoohPiglet,
                          method = "fdr")$Adjusted

PT
library(multcompView)
multcompLetters(PT,
                compare="<",
                threshold=0.05,
                Letters=letters)
```

pairwiseRobustTest *Pairwise two-sample robust tests*

Description

Performs pairwise two-sample robust tests across groups.

Usage

```
pairwiseRobustTest(formula = NULL, data = NULL, x = NULL, g = NULL,
                   est = "mom", nboot = 599, method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. $y \sim \text{group}$.
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
est	Estimate used for group comparisons. "onestep", "mom", "median", or "mean". See pb2gen for details.
nboot	The number of bootstrap samples.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to pb2gen .

Details

The input should include either formula and data; or x, and g.

The WRS2 package provides functions for robust estimation and hypothesis testing. This function invokes the pb2gen to make pairwise comparisons among groups.

The pairwiseRobustTest function can be used as a post-hoc method following an omnibus one-way anova with robust estimation.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/rcompanion/d_08a.html

See Also

[pairwiseRobustMatrix](#)

Examples

```

data(PoohPiglet)
PoohPiglet$Speaker = factor(PoohPiglet$Speaker,
                             levels = c("Pooh", "Tigger", "Piglet"))
PT = pairwiseRobustTest(Likert ~ Speaker,
                        data = PoohPiglet,
                        method = "fdr")
PT
cldList(comparison = PT$Comparison,
        p.value = PT$p.adjust,
        threshold = 0.05)

```

pairwiseSignMatrix *Pairwise sign tests for paired data with matrix output*

Description

Conducts pairwise sign tests across groups for paired data.

Usage

```

pairwiseSignMatrix(formula = NULL, data = NULL, x = NULL, g = NULL,
                   method = "fdr", ...)

```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to SIGN.test .

Details

The input should include either formula and data; or x, and g.

The two sample paired sign test compares medians among two groups with paired data. See http://rcompanion.org/handbook/F_07.html for further discussion of this test.

The `pairwiseSignTest` function can be used as a post-hoc method following an omnibus Friedman test. The matrix output can be converted to a compact letter display.

The function assumes that the data frame is already ordered by the blocking variable, so that the first observation of Group 1 is paired with the first observation of Group 2, and so on.

Value

A list consisting of: A matrix of p-values; the p-value adjustment method; a matrix of adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_10.html

See Also

[pairwiseSignTest](#)

Examples

```
data(BobBelcher)
friedman.test(Likert ~ Instructor | Rater,
             data = BobBelcher)
BobBelcher$Instructor = factor( BobBelcher$Instructor,
                               levels = c("Linda Belcher", "Louise Belcher",
                                           "Tina Belcher", "Bob Belcher",
                                           "Gene Belcher"))
PT = pairwiseSignMatrix(Likert ~ Instructor,
                       data = BobBelcher,
                       method = "fdr")$Adjusted
PT
library(multcompView)
multcompLetters(PT,
               compare="<",
               threshold=0.05,
               Letters=letters)
```

pairwiseSignTest

Pairwise sign tests for paired data

Description

Conducts pairwise sign tests across groups for paired data.

Usage

```
pairwiseSignTest(formula = NULL, data = NULL, x = NULL, g = NULL,  
  method = "fdr", ...)
```

Arguments

formula	A formula indicating the measurement variable and the grouping variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	The response variable as a vector.
g	The grouping variable as a vector.
method	The p-value adjustment method to use for multiple tests. See p.adjust .
...	Additional arguments passed to SIGN.test .

Details

The input should include either `formula` and `data`; or `x`, and `g`.

The two sample paired sign test compares medians among two groups with paired data. See http://rcompanion.org/handbook/F_07.html for further discussion of this test.

The `pairwiseSignTest` function can be used as a post-hoc method following an omnibus Friedman test.

The function assumes that the data frame is already ordered by the blocking variable, so that the first observation of Group 1 is paired with the first observation of Group 2, and so on.

Value

A dataframe of the groups being compared, the p-values, and the adjusted p-values.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the grouping variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_10.html

See Also

[pairwiseSignMatrix](#)

Examples

```

data(BobBelcher)
BobBelcher = BobBelcher[order(BobBelcher$Instructor, BobBelcher$Rater),]
friedman.test(Likert ~ Instructor | Rater,
              data = BobBelcher)
BobBelcher$Instructor = factor( BobBelcher$Instructor,
                                levels = c("Linda Belcher", "Louise Belcher",
                                             "Tina Belcher", "Bob Belcher",
                                             "Gene Belcher"))
PT = pairwiseSignTest(Likert ~ Instructor,
                     data = BobBelcher,
                     method = "fdr")

PT
cldList(comparison = PT$Comparison,
        p.value = PT$p.adjust,
        threshold = 0.05)

```

percentileTest	<i>Test of percentiles by permutation test</i>
----------------	--

Description

Conducts a permutation test to compare two groups for medians, percentiles, or proportion below a threshold value.

Usage

```

percentileTest(formula = NULL, data = NULL, x = NULL, y = NULL,
              test = "median", tau = 0.5, type = 7, threshold = NA,
              comparison = "<", r = 1000, digits = 4, progress = "TRUE")

```

Arguments

formula	A formula indicating the response variable and the independent variable. e.g. <code>y ~ group</code> .
data	The data frame to use.
x	If no formula is given, the response variable for one group.
y	The response variable for the other group.
test	The statistic to compare between groups. Can be "median", "percentile", "iqr", "proportion", "mean", or "variance".
tau	If "percentile" is chosen as the test, tau indicates the percentile to test. Expressed as a quantile. That is, 0.5 indicates a test for medians. 0.75 indicates a test for 75th percentiles.
type	The type value passed to the quantile function.

threshold	If "proportion" is chosen as the test, threshold indicates the value of the dependent variable to use as the threshold. For example, to test if there is a difference in the proportion of observations below \$10,000, threshold = 10000 would be used.
comparison	If "proportion" is chosen as the test, comparison indicates the inequality to use. Options are "<", "<=", ">", ">=", or, "=="
r	The number of replicates in the permutation test.
digits	The number of significant digits in the output.
progress	If TRUE, prints a dot for every 1 percent of progress while conducting the test.

Details

The function will test for a difference in medians, percentiles, interquartile ranges, proportion of observations above or below some threshold value, means, or variances between two groups by permutation test.

The input should include either formula and data; or x and y.

The function removes cases with NA in any of the variables.

If the independent variable has more than two groups, only the first two levels of the factor variable will be used.

The p-value returned is a two-sided test.

Value

A list of three data frames with the data used, a summary for each group, and the p-value from the test.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the independent variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/F_15.html

Examples

```
data(BrendonSmall)
percentileTest(Sodium ~ Instructor,
               data=BrendonSmall,
               test="median")

percentileTest(Sodium ~ Instructor,
               data=BrendonSmall,
```

```
        test="percentile",
        tau = 0.75)

percentileTest(Sodium ~ Instructor,
               data=BrendonSmall,
               test="proportion",
               threshold = 1300)
```

plotDensityHistogram *Histogram with a density curve*

Description

Produces a histogram for a vector of values and adds a density curve of the distribution.

Usage

```
plotDensityHistogram(x, prob = FALSE, col = "gray", main = "",
                    linecol = "black", lwd = 2, adjust = 1, bw = "nrd0",
                    kernel = "gaussian", ...)
```

Arguments

x	A vector of values.
prob	If FALSE, then counts are displayed in the histogram. If TRUE, then the density is shown.
col	The color of the histogram bars.
main	The title displayed for the plot.
linecol	The color of the line in the plot.
lwd	The width of the line in the plot.
adjust	Passed to density . A lower value makes the density plot smoother.
bw	Passed to density .
kernel	Passed to density .
...	Other arguments passed to hist .

Details

The function relies on the `hist` function. The density curve relies on the `density` function.

Value

Produces a plot. Returns nothing.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/C_04.html

See Also

[plotNormalHistogram](#) [plotNormalDensity](#)

Examples

```
### Plot of residuals from a model fit with lm
data(Catbus)
model = lm(Steps ~ Sex + Teacher,
           data = Catbus)
plotDensityHistogram(residuals(model))
```

plotNormalDensity *Density plot with a normal curve*

Description

Produces a density plot for a vector of values and adds a normal curve with the same mean and standard deviation. The plot can be used to quickly compare the distribution of data to a normal distribution.

Usage

```
plotNormalDensity(x, col1 = "white", col2 = "gray", col3 = "blue",
                 border = NA, main = "", lwd = 2, length = 1000, adjust = 1,
                 bw = "nrd0", kernel = "gaussian", ...)
```

Arguments

x	A vector of values.
col1	The color of the density plot. Usually not visible.
col2	The color of the density polygon.
col3	The color of the normal line.
border	The color of the border around the density polygon.
main	The title displayed for the plot.
lwd	The width of the line in the plot.
length	The number of points in the line in the plot.

adjust	Passed to density . A lower value makes the density plot smoother.
bw	Passed to density .
kernel	Passed to density .
...	Other arguments passed to plot .

Details

The function plots a polygon based on the [density](#) function. The normal curve has the same mean and standard deviation as the values in the vector.

Value

Produces a plot. Returns nothing.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/I_01.html

See Also

[plotNormalHistogram](#) [plotDensityHistogram](#)

Examples

```
### Plot of residuals from a model fit with lm
data(Catbus)
model = lm(Steps ~ Sex + Teacher,
           data = Catbus)
plotNormalDensity(residuals(model))
```

plotNormalHistogram *Histogram with a normal curve*

Description

Produces a histogram for a vector of values and adds a normal curve with the same mean and standard deviation. The plot can be used to quickly compare the distribution of data to a normal distribution.

Usage

```
plotNormalHistogram(x, prob = FALSE, col = "gray", main = "",
                    linecol = "blue", lwd = 2, length = 1000, ...)
```

Arguments

x	A vector of values.
prob	If FALSE, then counts are displayed in the histogram. If TRUE, then the density is shown.
col	The color of the histogram bars.
main	The title displayed for the plot.
linecol	The color of the line in the plot.
lwd	The width of the line in the plot.
length	The number of points in the line in the plot.
...	Other arguments passed to hist .

Details

The function relies on the `hist` function. The normal curve has the same mean and standard deviation as the values in the vector.

Value

Produces a plot. Returns nothing.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/I_01.html

See Also

[plotNormalDensity](#) [plotDensityHistogram](#)

Examples

```
### Plot of residuals from a model fit with lm
data(Catbus)
model = lm(Steps ~ Sex + Teacher,
           data = Catbus)
plotNormalHistogram(residuals(model))
```

plotPredy

*Plot a predicted line from a bivariate model***Description**

Plots the best fit line for a model with one y variable and one x variable, or with one y variable and polynomial x variables.

Usage

```
plotPredy(data, x, y, model, order = 1, x2 = NULL, x3 = NULL, x4 = NULL,
          x5 = NULL, pch = 16, xlab = "X", ylab = "Y", length = 1000,
          lty = 1, lwd = 2, col = "blue", type = NULL, ...)
```

Arguments

data	The name of the data frame.
x	The name of the x variable.
y	The name of the y variable.
model	The name of the model object.
order	If plotting a polynomial function, the order of the polynomial. Otherwise can be left as 1.
x2	If applicable, the name of the second order polynomial x variable.
x3	If applicable, the name of the third order polynomial x variable.
x4	If applicable, the name of the fourth order polynomial x variable.
x5	If applicable, the name of the fifth order polynomial x variable.
pch	The shape of the plotted data points.
xlab	The label for the x-axis.
ylab	The label for the y-axis.
length	The number of points used to draw the line.
lty	The style of the plotted line.
lwd	The width of the plotted line.
col	The col of the plotted line.
type	Passed to predict. Required for certain models.
...	Other arguments passed to plot.

Details

Any model for which predict() is defined can be used.

Value

Produces a plot. Returns nothing.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/I_10.html

Examples

```
### Plot of linear model fit with lm
data(BrendonSmall)
model = lm(Weight ~ Calories, data = BrendonSmall)
plotPredy(data = BrendonSmall,
           y = Weight,
           x = Calories,
           model = model,
           xlab = "Calories per day",
           ylab = "Weight in kilograms")

### Plot of polynomial model fit with lm
data(BrendonSmall)
BrendonSmall$Calories2 = BrendonSmall$Calories * BrendonSmall$Calories
model = lm(Sodium ~ Calories + Calories2, data = BrendonSmall)
plotPredy(data = BrendonSmall,
           y = Sodium,
           x = Calories,
           x2 = Calories2,
           model = model,
           order = 2,
           xlab = "Calories per day",
           ylab = "Sodium intake per day")

### Plot of quadratic plateau model fit with nls
data(BrendonSmall)
quadplat = function(x, a, b, clx) {
  ifelse(x < clx, a + b * x + (-0.5*b/clx) * x * x,
         a + b * clx + (-0.5*b/clx) * clx * clx)}
model = nls(Sodium ~ quadplat(Calories, a, b, clx),
            data = BrendonSmall,
            start = list(a = 519,
                        b = 0.359,
                        clx = 2304))
plotPredy(data = BrendonSmall,
           y = Sodium,
           x = Calories,
           model = model,
           xlab = "Calories per day",
           ylab = "Sodium intake per day")

### Logistic regression example requires type option
data(BullyHill)
Trials = cbind(BullyHill$Pass, BullyHill$Fail)
```



```
model.log = glm(Trials ~ Grade, data = BullyHill,
                family = binomial(link="logit"))
plotPredy(data = BullyHill,
           y     = Percent,
           x     = Grade,
           model = model.log,
           type  = "response",
           xlab  = "Grade",
           ylab  = "Proportion passing")
```

Pooh

Hypothetical data for paired ratings of Pooh Bear

Description

A data frame of Likert responses for instructor Pooh Bear for each of 10 respondents, paired before and after. Hypothetical data.

Usage

Pooh

Format

An object of class `data.frame` with 20 rows and 4 columns.

Source

http://rcompanion.org/handbook/F_06.html

PoohPiglet

Hypothetical data for ratings of Pooh, Piglet, and Tigger

Description

A data frame of Likert responses for instructors Pooh Bear, Piglet, and Tigger. Hypothetical data.

Usage

PoohPiglet

Format

An object of class `data.frame` with 30 rows and 2 columns.

Source

http://rcompanion.org/handbook/F_08.html

Religion	<i>Hypothetical data for change in religion after a caucusing event</i>
----------	---

Description

A matrix of paired counts for religion of people before and after an event. Hypothetical data.

Usage

Religion

Format

An object of class `matrix` with 4 rows and 4 columns.

Source

http://rcompanion.org/handbook/H_05.html

scheirerRayHare	<i>Scheirer Ray Hare test</i>
-----------------	-------------------------------

Description

Conducts Scheirer Ray Hare test.

Usage

```
scheirerRayHare(formula = NULL, data = NULL, y = NULL, x1 = NULL,
  x2 = NULL, tie.correct = TRUE, ss = TRUE, verbose = TRUE)
```

Arguments

formula	A formula indicating the response variable and two independent variables. e.g. $y \sim x1 + x2$.
data	The data frame to use.
y	If no formula is given, the response variable.
x1	The first independent variable.
x2	The second independent variable.
tie.correct	If "TRUE", applies a correction for ties in the response variable.
ss	If "TRUE", includes the sums of squares in the output.
verbose	If "TRUE", outputs statistics used in the analysis by direct print.

Details

The Scheirer Ray Hare test is a nonparametric test used for a two-way factorial experiment. It is described by Sokal and Rohlf (1995). At the time of writing, it does not appear to be a common or well-regarded test. It is sometimes recommended that the design should be balanced, and that there should be at least five observations for each cell in the interaction.

The input should include either formula and data; or y, x1, and x2.

The function removes cases with NA in any of the variables.

Value

A data frame of results similar to an anova table. Output from the verbose option is printed directly and not returned with the data frame.

Note

The parsing of the formula is simplistic. The first variable on the left side is used as the measurement variable. The first variable on the right side is used for the first independent variable. The second variable on the right side is used for the second independent variable.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

Sokal, R.R. and F.J. Rohlf. 1995. Biometry. 3rd ed. W.H. Freeman, New York.

http://rcompanion.org/handbook/F_14.html

Examples

```
### Example from Sokal and Rohlf, 1995.
Value = c(709,679,699,657,594,677,592,538,476,508,505,539)
Sex    = c(rep("Male",3), rep("Female",3), rep("Male",3), rep("Female",3))
Fat    = c(rep("Fresh", 6), rep("Rancid", 6))
Sokal = data.frame(Value, Sex, Fat)

scheirerRayHare(Value ~ Sex + Fat, data=Sokal)
```

transformTukey

Tukey's Ladder of Powers

Description

Conducts Tukey's Ladder of Powers on a vector of values to produce a more-normally distributed vector of values.

Usage

```
transformTukey(x, start = -10, end = 10, int = 0.025, plotit = TRUE,  
              verbose = FALSE, statistic = 1)
```

Arguments

x	A vector of values.
start	The starting value of lambda to try.
end	The ending value of lambda to try.
int	The interval between lambda values to try.
plotit	If TRUE, produces plots of Shapiro-Wilks W or Anderson-Darling A vs. lambda, a histogram of transformed values, and a quantile-quantile plot of transformed values.
verbose	If TRUE, prints extra output for Shapiro-Wilks W or Anderson-Darling A vs. lambda.
statistic	If 1, uses Shapiro-Wilks test. If 2, uses Anderson-Darling test.

Details

The function simply loops through lambda values from start to end at an interval of int.

The function then chooses the lambda which maximizes the Shapiro-Wilks W statistic or minimizes the Anderson-Darling A statistic.

It may be beneficial to add a constant to the input vector so that all values are positive. For left-skewed data, a (Constant - X) transformation may be helpful. Large values may need to be scaled.

Value

The transformed vector of values. The chosen lambda value is printed directly.

Author(s)

Salvatore Mangiafico, <mangiafico@njaes.rutgers.edu>

References

http://rcompanion.org/handbook/I_12.html

Examples

```
### Log-normal distribution example  
Conc = rlnorm(100)  
Conc.trans = transformTukey(Conc)
```

Index

*Topic **datasets**

- Anderson, 6
 - AndersonBias, 6
 - AndersonRainBarrel, 7
 - AndersonRainGarden, 7
 - BobBelcher, 8
 - Breakfast, 8
 - BrendonSmall, 9
 - BullyHill, 9
 - Catbus, 10
 - HayleySmith, 35
 - Monarchs, 36
 - Nurseries, 41
 - Pooh, 81
 - PoohPiglet, 81
 - Religion, 82
- accuracy, 4, 16, 17, 37
- Anderson, 6
- AndersonBias, 6
- AndersonRainBarrel, 7
- AndersonRainGarden, 7
- BobBelcher, 8
- boot.ci, 29, 31, 33
- Breakfast, 8
- BrendonSmall, 9
- BullyHill, 9
- Catbus, 10
- cateNelson, 4, 10
- chisq.test, 49, 50
- chisq_test, 52
- cldList, 12
- clm, 53, 58
- clmm, 55, 56
- cohenW, 14, 18
- compareGLM, 3, 5, 15, 17, 48
- compareLM, 3, 5, 16, 16, 48
- CramerV, 20
- cramerV, 14, 18
- cramerVFit, 14, 19
- density, 75, 77
- epsilonSquared, 20, 22
- fisher.test, 49, 50
- freemanTheta, 21, 22
- fullPTable, 23
- groupwiseCMH, 4, 24, 40, 44, 50, 51
- groupwiseGeometric, 25, 28, 30, 31, 33, 35
- groupwiseHuber, 26, 27, 30, 31, 33, 35
- groupwiseMean, 26, 28, 28, 31, 33, 35
- groupwiseMedian, 26, 28, 30, 30, 33, 35
- groupwisePercentile, 32, 60
- groupwiseSum, 34
- GTest, 49, 50
- HayleySmith, 35
- hist, 75, 78
- HuberM, 27
- independence_test, 61, 66
- MedianCI, 31
- Monarchs, 36
- mood.medtest, 45, 46
- multinomial.test, 40
- nagelkerke, 3, 5, 36, 39
- nagelkerkeHermite, 3, 37, 38
- nominalSymmetryTest, 25, 39, 44, 50, 51
- Nurseries, 41
- p.adjust, 40, 43, 45, 46, 49, 50, 52, 53, 55, 56, 58, 60, 61, 63, 64, 66, 67, 69, 70, 72
- pairwiseDifferences, 41
- pairwiseMcNemar, 4, 25, 40, 43, 50, 51

pairwiseMedianMatrix, [45](#), [47](#)
pairwiseMedianTest, [46](#), [46](#)
pairwiseModelAnova, [3](#), [16](#), [17](#), [48](#)
pairwiseNominalIndependence, [4](#), [25](#), [40](#),
[44](#), [49](#), [51](#), [52](#)
pairwiseNominalMatrix, [25](#), [40](#), [44](#), [50](#), [50](#)
pairwiseOrdinalIndependence, [51](#)
pairwiseOrdinalMatrix, [53](#), [58](#)
pairwiseOrdinalPairedMatrix, [54](#), [57](#)
pairwiseOrdinalPairedTest, [55](#), [56](#)
pairwiseOrdinalTest, [54](#), [57](#)
pairwisePercentileTest, [59](#)
pairwisePermutationMatrix, [61](#), [66](#)
pairwisePermutationSymmetry, [62](#), [65](#)
pairwisePermutationSymmetryMatrix, [63](#),
[64](#)
pairwisePermutationTest, [62](#), [65](#)
pairwiseRobustMatrix, [67](#), [69](#)
pairwiseRobustTest, [68](#), [68](#)
pairwiseSignMatrix, [70](#), [72](#)
pairwiseSignTest, [71](#), [71](#)
pb2gen, [67](#), [69](#)
percentileTest, [60](#), [73](#)
plot, [77](#)
plotDensityHistogram, [75](#), [77](#), [78](#)
plotNormalDensity, [76](#), [76](#), [78](#)
plotNormalHistogram, [3](#), [76](#), [77](#), [77](#)
plotPredy, [3](#), [79](#)
Pooh, [81](#)
PoohPiglet, [81](#)

rcompanion-package, [3](#)
Religion, [82](#)

scheirerRayHare, [82](#)
SIGN.test, [70](#), [72](#)
symmetry_test, [63](#), [64](#)

transformTukey, [83](#)

wilcox.test, [31](#)