Package 'lawstat'

July 22, 2025

Type Package

Title Tools for Biostatistics, Public Policy, and Law

Version 3.6

Date 2023-04-05

Depends R (>= 2.6.0)

Imports Kendall, mvtnorm, Rdpack

Suggests fBasics, tseries

License GPL (>= 2)

Description Statistical tests widely utilized in biostatistics, public pol-
icy, and law. Along with the well-known tests for equality of means and variances, random-
ness, and measures of relative variability, the package contains new robust tests of symme-
try, omnibus and directional tests of normality, and their graphical counterparts such as ro-
bust QQ plot, robust trend tests for variances, etc. All implemented tests and methods are illus-
trated by simulations and real-life examples from legal statistics, economics, and biostatistics.

RdMacros Rdpack

RoxygenNote 7.2.3

Encoding UTF-8

NeedsCompilation no

Author Joseph L. Gastwirth [aut],
Yulia R. Gel [aut, cre],
W. L. Wallace Hui [aut],
Vyacheslav Lyubchich [aut] (ORCID:
<https://orcid.org/0000-0001-7936-4285>),
Weiwen Miao [aut],
Kimihiro Noguchi [aut]

Maintainer Yulia R. Gel <gel.yulia@gmail.com>

Contents

---

bartels.test   *Ranked Version of von Neumann's Ratio Test for Randomness*

Description

Bartels (1982) test for randomness that is based on the ranked version of von Neumann's ratio (RVN). Users can choose whether to test against two-sided, negative, or positive correlation. NAs from the data are omitted.

Usage

```
bartels.test(
  y,
  alternative = c("two.sided", "positive.correlated", "negative.correlated")
)
```

Arguments

y	a numeric vector of data values.
alternative	a character string specifying the alternative hypothesis, must be one of `"two.sided"` (default), `"negative.correlated"`, or `"positive.correlated"`.

Value

A list of class `"htest"` with the following components:

statistic	the value of the standardized Bartels statistic.
parameter	RVN ratio.
p.value	the p-value for the test.
data.name	a character string giving the names of the data.
alternative	a character string describing the alternative hypothesis.

Author(s)

Kimihiro Noguchi, Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Bartels R (1982). "The rank version of von Neumann's ratio test for randomness." *Journal of the American Statistical Association*, **77**(377), 40–46. doi:10.1080/01621459.1982.10477764.

See Also

runs.test

Examples

```
## Simulate 100 observations from an autoregressive model of
## the first order AR(1)
y = arima.sim(n = 100, list(ar = c(0.5)))

## Test y for randomness
bartels.test(y)

## Sample Output
##
##        Bartels Test - Two sided
## data:  y
## Standardized Bartels Statistic -4.4929, RVN Ratio =
## 1.101, p-value = 7.024e-06
```

---

bias   *Prediction Errors ("Biases") of Surface Temperature Forecasts*

Description

Prediction errors of 48-hour ahead MM5 forecasts of surface temperature measured at 96 locations in the US Pacific Northwest on 3-January-2000. The prediction error, or "bias", is the difference between the forecasted and observed surface temperature. (MM5 is the fifth-generation Pennsylvania State University – National Center for Atmospheric Research Mesoscale Model.)

Usage

```
data(bias)
```

Format

A numeric vector of length 96.

Source

The data were kindly provided by the research group of Professor Clifford Mass in the Department of Atmospheric Sciences at the University of Washington. Detailed information about the Pacific Northwest prediction effort and the associated data archive can be found online at https://a.atmos.uw.edu/mm5rt/info.html and https://atmos.uw.edu/marka/pnw.html, respectively.

---

blackhire   *Hiring Data for Eight Professions and Two Races*

Description

Number of black and white candidates (hired or rejected) for eight professions (Gastwirth 1984).

Usage

```
data(blackhire)
```

Format

An array with 2 rows by 2 columns by 8 levels.

References

Gastwirth JL (1984). "Statistical methods for analyzing claims of employment discrimination." *ILR Review*, **38**(1), 75–86. doi:10.1177/001979398403800108.

---

brunner.munzel.test *Brunner–Munzel Test for Stochastic Equality*

Description

The Brunner–Munzel test for stochastic equality of two samples, which is also known as the Generalized Wilcoxon test. NAs from the data are omitted.

Usage

```
brunner.munzel.test(
  x,
  y,
  alternative = c("two.sided", "greater", "less"),
  alpha = 0.05
)
```

Arguments

x	the numeric vector of data values from the sample 1.
y	the numeric vector of data values from the sample 2.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less". User can specify just the initial letter.
alpha	significance level, default is 0.05 for 95% confidence interval.

Details

There exist discrepancies with Brunner and Munzel (2000) because there is a typo in the paper. The corrected version is in Neubert and Brunner (2007) (e.g., compare the estimates for the case study on pain scores). The current function follows Neubert and Brunner (2007).

Value

A list of class "htest" with the following components:

statistic	the Brunner–Munzel test statistic.
parameter	the degrees of freedom.
conf.int	the confidence interval.
p.value	the p-value of the test.
data.name	a character string giving the name of the data.
estimate	an estimate of the effect size, i.e., $P(X < Y) + 0.5P(X = Y)$.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao. This function was updated with the help of Dr. Ian Fellows.

References

Brunner E, Munzel U (2000). "The nonparametric Behrens–Fisher problem: asymptotic theory and a small-sample approximation." *Biometrical Journal*, **42**(1), 17–25.

Neubert K, Brunner E (2007). "A studentized permutation test for the non-parametric Behrens–Fisher problem." *Computational Statistics & Data Analysis*, **51**(10), 5192–5204. doi:10.1016/j.csda.2006.05.024.

See Also

wilcox.test, pwilcox

Examples

```
## Pain score on the third day after surgery for 14 patients under
## the treatment Y and 11 patients under the treatment N
## (see Brunner and Munzel, 2000; Neubert and Brunner, 2007).

Y <- c(1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 4, 1, 1)
N <- c(3, 3, 4, 3, 1, 2, 3, 1, 1, 5, 4)

brunner.munzel.test(Y, N)

##          Brunner-Munzel Test
## data: Y and N
## Brunner-Munzel Test Statistic = 3.1375,  df = 17.683, p-value = 0.005786
## 95 percent confidence interval:
##   0.5952169 0.9827052
## sample estimates:
## P(X<Y)+.5*P(X=Y)
##         0.788961
```

---

cd   *Coefficient of Dispersion – a Measure of Relative Variability*

Description

Measure of relative inequality (or relative variation) of the data. Coefficient of dispersion (CD) is the ratio of the mean absolute deviation from the median (MAAD) to the median of the data. NAs from the data are omitted. See Gastwirth (1988) and Bonett and Seier (2006).

Usage

```
cd(x)
```

Arguments

x   a numeric vector of data values.

Value

The coefficient of dispersion.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Bonett DG, Seier E (2006). "Confidence interval for a coefficient of dispersion in nonnormal distributions." *Biometrical Journal*, **48**(1), 144–148. doi:10.1002/bimj.200410148.

Gastwirth JL (1988). *Statistical Reasoning in Law and Public Policy: Statistical Concepts and Issues of Fairness*, volume 1. Academic Press, San Diego, CA.

See Also

gini.index, j.maad

Examples

```
## The Baker v. Carr Case: one-person-one-vote decision.
## Measure of Relative Inequality of Population data in 33 districts
## of the Tennessee Legislature in 1900 and 1972. See
## popdata (see Gastwirth, 1988).

data(popdata)
cd(popdata[,"pop1900"])
cd(popdata[,"pop1972"])
```

---

cmh.test   *The Cochran–Mantel–Haenszel Chi-square Test*

Description

The Cochran–Mantel–Haenszel (CMH) procedure tests homogeneity of population proportions after taking into account other factors. This procedure is widely used in law cases, for example, on equal employment and discrimination, and in biological and phamaceutical studies.

Usage

```
cmh.test(x)
```

Arguments

x   a numeric $2 \times 2 \times k$ array of data values.

Details

The test is based on the CMH procedure discussed by Gastwirth (1984). The data should be input in an array of 2 rows $\times$ 2 columns $\times$ k levels. The output includes the Mantel–Haenszel Estimate, the pooled Odd Ratio, and the Odd Ratio between the rows and columns at each level. The Chi-square test of significance tests if there is an interaction or association between rows and columns.

The null hypothesis is that the pooled Odd Ratio is equal to 1, i.e., there is no interaction between rows and columns. For more details see Gastwirth (1984).

The `cmh.test` can be viewed as a subset of `mantelhaen.test`, in the sense that `cmh.test` is for a 2 by 2 by k table without continuity correction, whereas `mantelhaen.test` allows for a larger table, and for a 2 by 2 by k table, it has an option of performing continuity correction. However, in view of Gastwirth (1984), continuity correction is not recommended as it tends to overestimate the p-value.

Value

A list of class `"htest"` containing the following components:

MH.ESTIMATE	the value of the Cochran–Mantel–Haenszel estimate.
OR	pooled Odd Ratio of the data.
ORK	vector of Odd Ratio of each level.
cmh	the test statistic.
df	degrees of freedom.
p.value	the p-value of the test.
method	type of the performed test.
data.name	a character string giving the name of the data.

Author(s)

Min Qin, Wallace W. Hui, Yulia R. Gel, Joseph L. Gastwirth

References

Gastwirth JL (1984). "Statistical methods for analyzing claims of employment discrimination." *ILR Review*, **38**(1), 75–86. doi:10.1177/001979398403800108.

See Also

mantelhaen.test

Examples

```
## Sample Salary Data
data(blackhire)
cmh.test(blackhire)
```

---

data1963   *Population Size and Number of Senators and Representatives in 1963*

Description

Number of senators and representatives and population size in 23 districts in the United States of America in 1963 (Gastwirth 1972).

Usage

```
data(data1963)
```

Format

A data frame with 23 observations on the following 3 variables:

pop1963   population in 1963;

sen1963   number of senators in the district in 1963;

rep1963   number of representatives in the district in 1963.

Source

Gastwirth (1972).

References

Gastwirth JL (1972). "The estimation of the Lorenz curve and Gini index." *The Review of Economics and Statistics*, **54**(3), 306–316.

---

gini.index   *Measures of Relative Variability – Gini Index*

Description

Gini index for measuring relative inequality (or relative variation) of the data (Gini 1912). NAs from the data are omitted.

Usage

```
gini.index(x)
```

Arguments

x   the input data.

Details

See also Gastwirth (1988).

Value

A list with the following components:

statistic   the Gini index.

parameter   the mean difference of the set of numbers.

data.name   a character string giving the name of the data.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Gastwirth JL (1988). *Statistical Reasoning in Law and Public Policy: Statistical Concepts and Issues of Fairness*, volume 1. Academic Press, San Diego, CA.

Gini C (1912). "Variabilita e mutabilita." *Reprinted in Memorie di Metodologica Statistica (Ed. Pizetti E. and Salvemini, T.), 1955, Rome: Libreria Eredi Virgilio Veschi.* English translation in Metron, 2005, 63(1): 3–38.

See Also

cd, j.maad, lorenz.curve

Examples

```
## The Baker v. Carr Case: one-person-one-vote decision.
## Measure of Relative Inequality of Population data in 33 districts
## of the Tennessee Legislature in 1900 and 1972. See
## popdata (see Gastwirth (1988)).
data(popdata)
gini.index(popdata[,"pop1900"])
gini.index(popdata[,"pop1972"])
```

---

j.maad   *MAAD Robust Standard Deviation*

Description

Compute average absolute deviation from the sample median, which is a consistent robust estimate of the population standard deviation for normally distribution data (Gastwirth 1982). NAs from the data are omitted.

Usage

```
j.maad(x)
```

Arguments

x   a numeric vector of data values.

Value

Robust standard deviation.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Gastwirth JL (1982). "Statistical properties of a measure of tax assessment uniformity." *Journal of Statistical Planning and Inference*, **6**(1), 1–12. doi:10.1016/03783758(82)900507.

See Also

cd, gini.index, rqq, rjb.test, sj.test

Examples

```
## Sample 100 observations from the standard normal distribution
x = rnorm(100)
j.maad(x)
```

---

laplace.test   *Goodness-of-fit Test Statistics for the Laplace Distribution*

Description

Goodness-of-fit test statistics A2 (Anderson–Darling), W2 (Cramer–von Mises), U2 (Watson), D (Kolmogorov–Smirnov), and V (Kuiper). By default, NAs are omitted. For the tables of critical values, see Stephens (1986) and Puig and Stephens (2000).

Usage

```
laplace.test(y)
```

Arguments

y   a numeric vector of data values.

Details

The function originally used `plaplace` function from R package VGAM (Yee 2019), however, to resolve dependencies between packages, the `plaplace` function was copied entirely to the current package under the name `VGAM_plaplace`.

Value

A list with the following numeric components:

A2	the Anderson–Darling statistic.
W2	the Cramer–von Mises statistic.
U2	the Watson statistic.
D	the Kolmogorov–Smirnov statistic.
V	the Kuiper statistic.

Author(s)

Kimihiro Noguchi, Yulia R. Gel

References

Puig P, Stephens MA (2000). "Tests of fit for the Laplace distribution, with applications." *Technometrics*, **42**(4), 417–424. doi:10.1080/00401706.2000.10485715.

Stephens MA (1986). "Tests for the Uniform Distribution." In D'Agostino RB, Stephens MA (eds.), *Goodness-of-fit Techniques*, volume 68 of *Statistics, textbooks and monographs*, chapter 8. Marcel Dekker, New York.

Yee T (2019). *VGAM: Vector Generalized Linear and Additive Models*. R package version 1.1-2, https://CRAN.R-project.org/package=VGAM.

Examples

```
## Differences in flood levels example taken from Puig and Stephens (2000)
y <- c(1.96,1.97,3.60,3.80,4.79,5.66,5.76,5.78,6.27,6.30,6.76,7.65,7.84,7.99,8.51,9.18,
    10.13,10.24,10.25,10.43,11.45,11.48,11.75,11.81,12.33,12.78,13.06,13.29,13.98,14.18,
     14.40,16.22,17.06)
laplace.test(y)$D
## [1] 0.9177726
## The critical value at the 0.05 significance level is approximately 0.906.
## Thus, the null hypothesis should be rejected at the 0.05 level.
```

---

levene.test   *Levene's Test of Equality of Variances*

Description

Tests equality of the k population variances.

Usage

```
levene.test(
  y,
  group,
  location = c("median", "mean", "trim.mean"),
  trim.alpha = 0.25,
  bootstrap = FALSE,
  num.bootstrap = 1000,
  kruskal.test = FALSE,
 correction.method = c("none", "correction.factor", "zero.removal", "zero.correction")
)
```

Arguments

y	a numeric vector of data values.
group	factor of the data.
location	the default option is "median" corresponding to the robust Brown–Forsythe Levene-type procedure (Brown and Forsythe 1974); "mean" corresponds to the classical Levene's procedure (Levene 1960), and "trim.mean" corresponds to the robust Levene-type procedure using the group trimmed means.
trim.alpha	the fraction (0 to 0.5) of observations to be trimmed from each end of x before the mean is computed.
bootstrap	a logical value identifying whether to implement bootstrap. The default is FALSE, i.e., no bootstrap; if set to TRUE, the bootstrap method described in Lim and Loh (1996) for Levene's test is applied.
num.bootstrap	number of bootstrap samples to be drawn when the bootstrap argument is set to TRUE. The default value is 1000.
kruskal.test	logical value indentifying whether to use the Kruskal–Wallis statistic. The default option is FALSE, i.e., the usual ANOVA statistic is used.
correction.method	
	procedures to make the test more robust; the default option is "none"; "correction.factor" applies the correction factor described by O'Brien (1978) and Keyes and Levy (1997); "zero.removal" performs the structural zero removal method by Hines and Hines (2000); "zero.correction" performs a combination of the O'Brien's correction factor and the Hines–Hines structural zero removal method (Noguchi and Gel 2010). Note that the options "zero.removal" and "zero.correction" are only applicable when the location is set to "median", otherwise, "none" is applied.

Details

The test statistic is based on the classical Levene's procedure (using the group means), the modified Brown–Forsythe Levene-type procedure (using the group medians), or the modified Levene-type procedure (using the group trimmed means). More robust versions of the test using the correction factor or structural zero removal method are also available. Two options for calculating critical values, namely, approximated and bootstrapped, are available. By default, NAs are omitted from the data.

Value

A list of class `"htest"` with the following components:

`statistic`	the value of the test statistic.
`p.value`	the p-value of the test.
`method`	type of test performed.
`data.name`	a character string giving the name of the data.
`non.bootstrap.p.value`	
	the p-value of the test without bootstrap method; i.e. the p-value using the approximated critical value.

Note

Instead of the ANOVA statistic suggested by Levene, the Kruskal–Wallis ANOVA may also be applied using this function (see the parameter `kruskal.test`).

Modified from a response posted by Brian Ripley to the R-help e-mail list.

Author(s)

Kimihiro Noguchi, W. Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Brown MB, Forsythe AB (1974). "Robust tests for the equality of variances." *Journal of the American Statistical Association*, **69**(346), 364–367. doi:10.1080/01621459.1974.10482955.

Hines WGS, Hines RJO (2000). "Increased power with modified forms of the Levene (Med) test for heterogeneity of variance." *Biometrics*, **56**(2), 451–454. doi:10.1111/j.0006341X.2000.00451.x.

Keyes TK, Levy MS (1997). "Analysis of Levene's test under design imbalance." *Journal of Educational and Behavioral Statistics*, **22**(2), 227–236. doi:10.3102/10769986022002227.

Levene H (1960). "Robust Tests for Equality of Variances." In Olkin I, others (eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Palo Alto, CA.

Lim T, Loh W (1996). "A comparison of tests of equality of variances." *Computational Statistics & Data Analysis*, **22**(3), 287–301. doi:10.1016/01679473(95)000542.

Noguchi K, Gel YR (2010). "Combination of Levene-type tests and a finite-intersection method for testing equality of variances against ordered alternatives." *Journal of Nonparametric Statistics*, **22**(7), 897–913. doi:10.1080/10485251003698505.

O'Brien RG (1978). "Robust techniques for testing heterogeneity of variance effects in factorial designs." *Psychometrika*, **43**(3), 327–342. doi:10.1007/BF02293643.

See Also

neuhauser.hothorn.test, lnested.test, ltrend.test, mma.test, robust.mmm.test

Examples

```
data(pot)
levene.test(pot[,"obs"], pot[,"type"],
            location = "median", correction.method = "zero.correction")

## Bootstrap version of the test. The calculation may take up a few minutes
## depending on the number of bootstrap sampling.
levene.test(pot[,"obs"], pot[,"type"],
            location = "median", correction.method = "zero.correction",
            bootstrap = TRUE, num.bootstrap = 500)
```

---

lnested.test	*Test for a Monotonic Trend in Variances*

Description

The test statistic is based on the finite intersection approach.

Usage

```
lnested.test(
  y,
  group,
  location = c("median", "mean", "trim.mean"),
  tail = c("right", "left", "both"),
  trim.alpha = 0.25,
  bootstrap = FALSE,
  num.bootstrap = 1000,
 correction.method = c("none", "correction.factor", "zero.removal", "zero.correction"),
  correlation.method = c("pearson", "kendall", "spearman")
)
```

Arguments

y	a numeric vector of data values.
group	factor of the data.
location	the default option is `"median"` corresponding to the robust Brown–Forsythe Levene-type procedure (Brown and Forsythe 1974); `"mean"` corresponds to the classical Levene's procedure (Levene 1960), and `"trim.mean"` corresponds to the robust Levene-type procedure using the group trimmed means.
tail	the default option is `"right"`, corresponding to an increasing trend in variances as the one-sided alternative; `"left"` corresponds to a decreasing trend in variances, and `"both"` corresponds to any (increasing or decreasing) monotonic trend in variances as the two-sided alternative.
trim.alpha	the fraction (0 to 0.5) of observations to be trimmed from each end of x before the mean is computed.
bootstrap	a logical value identifying whether to implement bootstrap. The default is `FALSE`, i.e., no bootstrap; if set to `TRUE`, the bootstrap method described in Lim and Loh (1996) for Levene's test is applied.
num.bootstrap	number of bootstrap samples to be drawn when the `bootstrap` argument is set to `TRUE`. The default value is 1000.
correction.method	
	procedures to make the test more robust; the default option is `"none"`; `"correction.factor"` applies the correction factor described by O'Brien (1978) and Keyes and Levy (1997); `"zero.removal"` performs the structural zero removal method by Hines and Hines (2000); `"zero.correction"` performs a combination of the O'Brien's correction factor and the Hines–Hines structural zero removal method (Noguchi and Gel 2010). Note that the options `"zero.removal"` and `"zero.correction"` are only applicable when the location is set to `"median"`, otherwise, `"none"` is applied.
correlation.method	
	measures of correlation; the default option is `"pearson"`, the linear correlation coefficient that is equivalent to the t-test; nonparametric measures of correlation such as `"kendall"` (Kendall's tau) or `"spearman"` (Spearman's rho) may also be chosen.

Details

The test statistic is based on the classical Levene's procedure (using the group means), the modified Brown–Forsythe Levene-type procedure (using the group medians), or the modified Levene-type procedure (using the group trimmed means). More robust versions of the test using the correction factor or structural zero removal method are also available. Two options for calculating critical values, namely, approximated and bootstrapped, are available. By default, NAs are omitted from the data.

Value

A list with the following elements:

T	the statistic and p-value of the test based on the Tippett p-value combination.

F	the statistic and p-value of the test based on the Fisher p-value combination.
N	the statistic and p-value of the test based on the Liptak p-value combination.
L	the statistic and p-value of the test based on the Mudholkar–George p-value combination.

Each of the list elements is a list of class `"htest"` with the following elements:

statistic	the value of the test statistic expressed in terms of correlation (Pearson, Kendall, or Spearman).
p.value	the p-value of the test.
method	type of test performed.
data.name	a character string giving the name of the data.
non.bootstrap.statistic	
	the statistic of the test without bootstrap method.
non.bootstrap.p.value	
	the p-value of the test without bootstrap method.

Author(s)

Kimihiro Noguchi, W. Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Brown MB, Forsythe AB (1974). "Robust tests for the equality of variances." *Journal of the American Statistical Association*, **69**(346), 364–367. doi:10.1080/01621459.1974.10482955.

Hines WGS, Hines RJO (2000). "Increased power with modified forms of the Levene (Med) test for heterogeneity of variance." *Biometrics*, **56**(2), 451–454. doi:10.1111/j.0006341X.2000.00451.x.

Keyes TK, Levy MS (1997). "Analysis of Levene's test under design imbalance." *Journal of Educational and Behavioral Statistics*, **22**(2), 227–236. doi:10.3102/10769986022002227.

Levene H (1960). "Robust Tests for Equality of Variances." In Olkin I, others (eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Palo Alto, CA.

Lim T, Loh W (1996). "A comparison of tests of equality of variances." *Computational Statistics & Data Analysis*, **22**(3), 287–301. doi:10.1016/01679473(95)000542.

Noguchi K, Gel YR (2010). "Combination of Levene-type tests and a finite-intersection method for testing equality of variances against ordered alternatives." *Journal of Nonparametric Statistics*, **22**(7), 897–913. doi:10.1080/10485251003698505.

O'Brien RG (1978). "Robust techniques for testing heterogeneity of variance effects in factorial designs." *Psychometrika*, **43**(3), 327–342. doi:10.1007/BF02293643.

See Also

levene.test, ltrend.test, mma.test, neuhauser.hothorn.test, robust.mmm.test

Examples

```
data(pot)
lnested.test(pot[,"obs"], pot[, "type"], location = "median", tail = "left",
             correction.method = "zero.correction")$N

lnested.test(pot[, "obs"], pot[, "type"], location = "median", tail = "left",
             correction.method = "zero.correction",
             bootstrap = TRUE, num.bootstrap = 500)$N
```

---

lorenz.curve   *Lorenz Curve*

Description

Plots the Lorenz curve that is a graphical representation of the cumulative distribution function. The user can choose between the Lorenz curve with single (default) or multiple weighting of data, for example, taking into account for single or multiple legislature representatives (Gastwirth 1972).

Usage

```
lorenz.curve(
  data,
  weight = NULL,
  mul = FALSE,
  plot.it = TRUE,
  main = NULL,
  xlab = NULL,
  ylab = NULL,
  xlim = c(0, 1),
  ylim = c(0, 1),
  ...
)
```

Arguments

data   input data. If the argument is an array, a matrix, a data.frame, or a list with two or more columns, then the first column will be treated as a data vector, and the second column to be treated as a weight vector. A separate weight vector is then ignored and not required. If the argument is a single column vector, then a user must enter a separate single-column weight vector. NAs or character are not allowed.

weight   one-column vector contains factors of single or multiple weights. Ignored if included in the data argument. NAs or character are not allowed.

mul   logical value indicates whether the Lorenz curve with multiple weight is to be plotted. Default is FALSE, i.e., single.

plot.it	logical value indicates whether the Lorenz curve should be plotted. Default is TRUE, i.e., to plot.
main	title of Lorenz curve. Only required if user wants to override the default value.
xlab	label of x-axis. Only required if user wants to override the default value.
ylab	label of y-axis. Only required if user wants to override the default value.
xlim	plotting range of x-axis. Only required if user wants to override the default value.
ylim	plotting range of y-axis. Only required if user wants to override the default value.
...	other graphical parameters to be passed to the [plot](#) function.

Details

The input data should be a data frame with 2 columns. The first column will be treated as data vector, and the second column to be treated as a weight vector. Alternatively, data and weights can be entered as separate one-column vectors.

Value

A Lorenz curve plot with x-axis being the culmulative fraction of the data argument, and y-axis being the culmulative fraction of the weight argument. In the legend to the plot, the following values are reported:

RMD	relative mean deviation of the input data.
GI	the Gini index of the input data.
L(1/2)	median of the culmulative fraction sum of the data.

Author(s)

Man Jin, Wallace W. Hui, Yulia R. Gel, Joseph L. Gastwirth

References

Gastwirth JL (1972). "The estimation of the Lorenz curve and Gini index." *The Review of Economics and Statistics*, **54**(3), 306–316.

See Also

[gini.index](#)

Examples

```
## Data on: number of senators (second column) and
## representatives (third column) relative to population size (first column) in 1963
## First column is treated as the data argument.
data(data1963)

## Single weight Lorenz Curve using number of senators as weight argument.
```

```
lorenz.curve(data1963)

## Multiple weight Lorenz Curve using number of senators as weight argument.
lorenz.curve(data1963, mul = TRUE)

## Multiple weight Lorenz Curve using number of representatives
## as weight argument.
lorenz.curve(data1963[, "pop1963"], data1963[, "rep1963"], mul = TRUE)
```

---

ltrend.test   *Test for a Linear Trend in Variances*

Description

Test for a linear trend in variances.

Usage

```
ltrend.test(
  y,
  group,
  score = NULL,
  location = c("median", "mean", "trim.mean"),
  tail = c("right", "left", "both"),
  trim.alpha = 0.25,
  bootstrap = FALSE,
  num.bootstrap = 1000,
 correction.method = c("none", "correction.factor", "zero.removal", "zero.correction"),
  correlation.method = c("pearson", "kendall", "spearman")
)
```

Arguments

y	a numeric vector of data values.
group	factor of the data.
score	weights to be used in testing an increasing/decreasing trend in group variances, score coincides by default with group; it can be chosen as a linear, quadratic or any other monotone function.
location	the default option is "median" corresponding to the robust Brown–Forsythe Levene-type procedure (Brown and Forsythe 1974); "mean" corresponds to the classical Levene's procedure (Levene 1960), and "trim.mean" corresponds to the robust Levene-type procedure using the group trimmed means.
tail	the default option is "right", corresponding to an increasing trend in variances as the one-sided alternative; "left" corresponds to a decreasing trend in variances, and "both" corresponds to any (increasing or decreasing) monotonic trend in variances as the two-sided alternative.

`trim.alpha`	the fraction (0 to 0.5) of observations to be trimmed from each end of `x` before the mean is computed.
`bootstrap`	a logical value identifying whether to implement bootstrap. The default is `FALSE`, i.e., no bootstrap; if set to `TRUE`, the bootstrap method described in Lim and Loh (1996) for Levene's test is applied.
`num.bootstrap`	number of bootstrap samples to be drawn when the `bootstrap` argument is set to `TRUE`. The default value is 1000.
`correction.method`	
	procedures to make the test more robust; the default option is `"none"`; `"correction.factor"` applies the correction factor described by O'Brien (1978) and Keyes and Levy (1997); `"zero.removal"` performs the structural zero removal method by Hines and Hines (2000); `"zero.correction"` performs a combination of the O'Brien's correction factor and the Hines–Hines structural zero removal method (Noguchi and Gel 2010). Note that the options `"zero.removal"` and `"zero.correction"` are only applicable when the location is set to `"median"`, otherwise, `"none"` is applied.
`correlation.method`	
	measures of correlation; the default option is `"pearson"`, the linear correlation coefficient that is equivalent to the t-test; nonparametric measures of correlation such as `"kendall"` (Kendall's tau) or `"spearman"` (Spearman's rho) may also be chosen.

Details

The test statistic is based on the classical Levene's procedure (using the group means), the modified Brown–Forsythe Levene-type procedure (using the group medians), or the modified Levene-type procedure (using the group trimmed means). More robust versions of the test using the correction factor or structural zero removal method are also available. Two options for calculating critical values, namely, approximated and bootstrapped, are available. By default, NAs are omitted from the data.

Value

A list of class `"htest"` containing the following components:

`statistic`	the value of the test statistic expressed in terms of correlation (Pearson, Kendall, or Spearman).
`p.value`	the p-value of the test.
`method`	type of test performed.
`data.name`	a character string giving the name of the data.
`t.statistic`	the value of the test statistic from Student's t-test.
`non.bootstrap.p.value`	
	the p-value of the test without bootstrap method.
`log.p.value`	the log of the p-value
`log.q.value`	the log of the (one minus the p-value).

Author(s)

Kimihiro Noguchi, W. Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Brown MB, Forsythe AB (1974). "Robust tests for the equality of variances." *Journal of the American Statistical Association*, **69**(346), 364–367. doi:10.1080/01621459.1974.10482955.

Hines WGS, Hines RJO (2000). "Increased power with modified forms of the Levene (Med) test for heterogeneity of variance." *Biometrics*, **56**(2), 451–454. doi:10.1111/j.0006341X.2000.00451.x.

Keyes TK, Levy MS (1997). "Analysis of Levene's test under design imbalance." *Journal of Educational and Behavioral Statistics*, **22**(2), 227–236. doi:10.3102/10769986022002227.

Levene H (1960). "Robust Tests for Equality of Variances." In Olkin I, others (eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Palo Alto, CA.

Lim T, Loh W (1996). "A comparison of tests of equality of variances." *Computational Statistics & Data Analysis*, **22**(3), 287–301. doi:10.1016/01679473(95)000542.

Noguchi K, Gel YR (2010). "Combination of Levene-type tests and a finite-intersection method for testing equality of variances against ordered alternatives." *Journal of Nonparametric Statistics*, **22**(7), 897–913. doi:10.1080/10485251003698505.

O'Brien RG (1978). "Robust techniques for testing heterogeneity of variance effects in factorial designs." *Psychometrika*, **43**(3), 327–342. doi:10.1007/BF02293643.

See Also

neuhauser.hothorn.test, levene.test, lnested.test, mma.test, robust.mmm.test

Examples

```
data(pot)
ltrend.test(pot[, "obs"], pot[, "type"], location = "median", tail = "left",
            correction.method = "zero.correction")

## Bootstrap version of the test. The calculation may take up a few minutes
## depending on the number of bootstrap samples.
ltrend.test(pot[, "obs"], pot[, "type"], location = "median", tail = "left",
             correction.method = "zero.correction",
             bootstrap = TRUE, num.bootstrap = 500)
```

---

michigan	*Dioxin Levels for Counties in the Upper Peninsula of Michigan*

Description

Data contains 16 observations of dioxin levels for counties in the Upper Peninsula of Michigan.

Usage

```
data(michigan)
```

Format

A numeric vector of length 16.

Source

The Environmental Protection Agency (EPA) of the State of Michigan.

---

mma.test	*Mudholkar–McDermott–Aumont Test for Ordered Variances for Normal Samples*

Description

Test for a monotonic trend in variances for normal samples. The test statistic is based on a combination of the finite intersection approach and the classical F (variance ratio) test (Mudholkar et al. 1993). By default, NAs are omitted.

Usage

```
mma.test(y, group, tail = c("right", "left", "both"))
```

Arguments

y	a numeric vector of data values.
group	factor of the data.
tail	the default option is "right", corresponding to an increasing trend in variances as the one-sided alternative; "left" corresponds to a decreasing trend in variances, and "both" corresponds to any (increasing or decreasing) monotonic trend in variances as the two-sided alternative.

Value

A list with the following components:

T	the statistic and p-value of the test based on the Tippett p-value combination.
F	the statistic and p-value of the test based on the Fisher p-value combination.
N	the statistic and p-value of the test based on the Liptak p-value combination.
L	the statistic and p-value of the test based on the Mudholkar–George p-value combination.

Each of the list elements is a list of class `"htest"` with the following elements:

statistic	the value of the test statistic.
p.value	the p-value of the test.
method	type of test performed.
data.name	a character string giving the name of the data.

Author(s)

Kimihiro Noguchi, Yulia R. Gel

References

Mudholkar GS, McDermott MP, Aumont J (1993). "Testing homogeneity of ordered variances." *Metrika*, **40**(1), 271–281. doi:10.1007/BF02613691.

See Also

neuhauser.hothorn.test, levene.test, lnested.test, ltrend.test, robust.mmm.test

Examples

```
data(pot)
mma.test(pot[, "obs"], pot[, "type"], tail = "left")$N
```

---

neuhauser.hothorn.test

Neuhauser–Hothorn Double Contrast Test for a Monotonic Trend in Variances

---

Description

The test statistic suggested by Neuhauser and Hothorn (2000).

Usage

```
neuhauser.hothorn.test(
  y,
  group,
  location = c("median", "mean", "trim.mean"),
  tail = c("right", "left", "both"),
  trim.alpha = 0.25,
  bootstrap = FALSE,
  num.bootstrap = 1000,
  correction.method = c("none", "correction.factor", "zero.removal", "zero.correction")
)
```

Arguments

y   a numeric vector of data values.

group   factor of the data.

location   the default option is "median" corresponding to the robust Brown–Forsythe
                    Levene-type procedure (Brown and Forsythe 1974); "mean" corresponds to the
                    classical Levene's procedure (Levene 1960), and "trim.mean" corresponds to
                    the robust Levene-type procedure using the group trimmed means.

tail   the default option is "right", corresponding to an increasing trend in variances
                    as the one-sided alternative; "left" corresponds to a decreasing trend in vari-
                    ances, and "both" corresponds to any (increasing or decreasing) monotonic
                    trend in variances as the two-sided alternative.

trim.alpha   the fraction (0 to 0.5) of observations to be trimmed from each end of x before
                    the mean is computed.

bootstrap   a logical value identifying whether to implement bootstrap. The default is FALSE,
                    i.e., no bootstrap; if set to TRUE, the bootstrap method described in Lim and Loh
                    (1996) for Levene's test is applied.

num.bootstrap   number of bootstrap samples to be drawn when the bootstrap argument is set
                    to TRUE. The default value is 1000.

correction.method
                    procedures to make the test more robust; the default option is "none"; "correction.factor"
                    applies the correction factor described by O'Brien (1978) and Keyes and Levy
                    (1997); "zero.removal" performs the structural zero removal method by Hines
                    and Hines (2000); "zero.correction" performs a combination of the O'Brien's
                    correction factor and the Hines–Hines structural zero removal method (Noguchi
                    and Gel 2010). Note that the options "zero.removal" and "zero.correction"
                    are only applicable when the location is set to "median", otherwise, "none" is
                    applied.

Details

The test statistic is based on the classical Levene's procedure (using the group means), the modified
Brown–Forsythe Levene-type procedure (using the group medians), or the modified Levene-type
procedure (using the group trimmed means). More robust versions of the test using the correction
factor or structural zero removal method are also available. Two options for calculating critical

values, namely, approximated and bootstrapped, are available. By default, NAs are omitted from the
data.

Value

A list of class `"htest"` with the following components:

statistic   the value of the test statistic.

p.value   the p-value of the test.

method   type of test performed.

data.name   a character string giving the name of the data.

non.bootstrap.p.value

          the p-value of the test without bootstrap method.

Author(s)

Kimihiro Noguchi, Yulia R. Gel

References

Brown MB, Forsythe AB (1974). "Robust tests for the equality of variances." *Journal of the American Statistical Association*, **69**(346), 364–367. doi:10.1080/01621459.1974.10482955.

Hines WGS, Hines RJO (2000). "Increased power with modified forms of the Levene (Med) test for heterogeneity of variance." *Biometrics*, **56**(2), 451–454. doi:10.1111/j.0006341X.2000.00451.x.

Keyes TK, Levy MS (1997). "Analysis of Levene's test under design imbalance." *Journal of Educational and Behavioral Statistics*, **22**(2), 227–236. doi:10.3102/10769986022002227.

Levene H (1960). "Robust Tests for Equality of Variances." In Olkin I, others (eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Palo Alto, CA.

Lim T, Loh W (1996). "A comparison of tests of equality of variances." *Computational Statistics & Data Analysis*, **22**(3), 287–301. doi:10.1016/01679473(95)000542.

Neuhauser M, Hothorn LA (2000). "Parametric location-scale and scale trend tests based on Levene's transformation." *Computational Statistics & Data Analysis*, **33**(2), 189–200. doi:10.1016/S01679473(99)000511.

Noguchi K, Gel YR (2010). "Combination of Levene-type tests and a finite-intersection method for testing equality of variances against ordered alternatives." *Journal of Nonparametric Statistics*, **22**(7), 897–913. doi:10.1080/10485251003698505.

O'Brien RG (1978). "Robust techniques for testing heterogeneity of variance effects in factorial designs." *Psychometrika*, **43**(3), 327–342. doi:10.1007/BF02293643.

See Also

levene.test, lnested.test, ltrend.test, mma.test, robust.mmm.test

Examples

```
data(pot)
neuhauser.hothorn.test(pot[, "obs"], pot[, "type"], location = "median",
                       tail = "left", correction.method = "zero.correction")

## Bootstrap version of the test. The calculation may take up a few minutes
## depending on the number of bootstrap sampling.
neuhauser.hothorn.test(pot[, "obs"], pot[, "type"], location = "median",
                       tail = "left", correction.method = "zero.correction",
                       bootstrap = TRUE, num.bootstrap = 500)
```

---

nig.parameter	*Generate Parameters for the Normal Inverse Gaussian (NIG) Distribution*

Description

Produce four parameters, alpha (tail heavyness), beta (asymmetry), delta (scale), and mu (location) from the four variables: mean, variance, kurtosis, and skewness.

Usage

```
nig.parameter(
  mean = mean,
  variance = variance,
  kurtosis = kurtosis,
  skewness = skewness
)
```

Arguments

mean	mean of the NIG distribution.
variance	variance of the NIG distribution.
kurtosis	excess kurtosis of the NIG distribution.
skewness	skewness of the NIG distribution.

Details

The parameters are generated with three conditions: 1) $3 \times kurtosis > 5 \times skewness^2$; 2) $skewness > 0$, and 3) $variance > 0$. See Atkinson (1982), Barndorff-Nielsen and Blaesild (1983), and Noguchi and Gel (2010).

Value

A list with the following numeric components:

alpha   tail-heavyness parameter of the NIG distribution.

beta   asymmetry parameter of the NIG distribution.

delta   scale parameter of the NIG distribution.

mu   location parameter of the NIG distribution.

Author(s)

Kimihiro Noguchi, Yulia R. Gel

References

Atkinson AC (1982). "The simulation of generalized inverse Gaussian and hyperbolic random variables." *SIAM Journal on Scientific and Statistical Computing*, **3**(4), 502–515. doi:10.1137/0903033.

Barndorff-Nielsen OE, Blaesild P (1983). "Hyperbolic distributions." In Johnson NL, Kotz S, Read CB (eds.), *Encyclopedia of Statistical Sciences*, 700–707. John Wiley & Sons Ltd, New York.

Noguchi K, Gel YR (2010). "Combination of Levene-type tests and a finite-intersection method for testing equality of variances against ordered alternatives." *Journal of Nonparametric Statistics*, **22**(7), 897–913. doi:10.1080/10485251003698505.

See Also

rnig

Examples

```
library(fBasics)
test <- nig.parameter(0, 2, 5, 1)
random <- rnig(1000000, alpha = test$alpha, beta = test$beta,
               mu = test$mu, delta = test$delta)
mean(random)
var(random)
kurtosis(random)
skewness(random)
```

popdata	*Population Size of 33 Districts of the Tennessee Legislature in 1900, 1960, and 1972*

Description

The Baker v. Carr Case: one-person-one-vote decision. Measure of Relative Inequality of Population data in 33 districts of the Tennessee Legislature in 1900, 1960, and 1972 (Gastwirth 1988).

Usage

```
data(popdata)
```

Format

A data frame with 33 observations on the following 3 numeric variables:

pop1900 population data in 1900

pop1960 population data in 1960

pop1972 population data in 1972

Source

Gastwirth (1988).

References

Gastwirth JL (1988). *Statistical Reasoning in Law and Public Policy: Statistical Concepts and Issues of Fairness*, volume 1. Academic Press, San Diego, CA.

pot	*Apertures of Chupa Pots from Three Philippine Communities*

Description

The apertures of the chupa pots from three Philippine locations: Dalupa (`ApDl`), Dangtalan (`ApDg`), and Paradijon (`ApP`).

Usage

```
data(pot)
```

Format

A data frame with 343 observations of 2 variables: obs (integer values of observed apertures) and locations (factor with 3 levels).

Details

Archaeologists are concerned with the effect that increasing economic activity had on older civilizations. Economic growth and its related economic specialization led to the "standardization hypothesis", i.e., increased production of an item would lead to its becoming more uniform. Kvamme et al. (1996) focused on earthenware, chupa-pots from three Philippine communities that differ in the way they organize ceramic production. In Dangtalan, pottery is primarily made for household use; in Dalupa there is a non-market barter economy where potters exchange their works. In the village of Paradijon, near the provincial capital, full-time pottery specialists sell their output to shopkeepers for sale to the general public.

Source

The data are kindly provided by Professor Kvamme (Kvamme et al. 1996).

References

Kvamme KL, Stark MT, Longacre WA (1996). "Alternative procedures for assessing standardization in ceramic assemblages." *American Antiquity*, **61**(1), 116–126. doi:10.2307/282306.

---

rjb.test   *Test of Normailty – Robust Jarque–Bera Test*

Description

The robust and classical Jarque–Bera tests of normality.

Usage

```
rjb.test(
  x,
  option = c("RJB", "JB"),
  crit.values = c("chisq.approximation", "empirical"),
  N = 0
)
```

Arguments

x	a numeric vector of data values.
option	the choice of whether to perform the robust test, "RJB" (default) or classic test, "JB".
crit.values	a character string specifying how the critical values should be obtained: approximated by the Chi-square distribution (default) or empirically.
N	number of Monte Carlo simulations for the empirical critical values.

Details

The test is based on a joint statistic using skewness and kurtosis coefficients. The Robust Jarque–Bera (RJB) is the robust version of the Jarque–Bera (JB) test of normality. The RJB (default option) utilizes the robust standard deviation (specifically, the Average Absolute Deviation from the Median; MAAD) to estimate sample kurtosis and skewness. For more details, see Gel and Gastwirth (2008). Users can also choose to perform the classical Jarque–Bera test (Jarque and Bera 1980).

Value

A list of class `"htest"` with the following components:

statistic	the value of the test statistic.
parameter	the degrees of freedom.
p.value	the p-value of the test.
method	type of test was performed.
data.name	a character string giving the name of the data.

Note

Modified from `jarque.bera.test` (tseries package).

Author(s)

W. Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Gel YR, Gastwirth JL (2008). "A robust modification of the Jarque–Bera test of normality." *Economics Letters*, **99**(1), 30–32. doi:10.1016/j.econlet.2007.05.022.

Jarque CM, Bera AK (1980). "Efficient tests for normality, homoscedasticity and serial independence of regression residuals." *Economics Letters*, **6**(3), 255–259. doi:10.1016/01651765(80)90024-5.

See Also

sj.test, rqq, jarque.bera.test

Examples

```
## Normally distributed data
x = rnorm(100)
rjb.test(x)

## Using zuni data
data(zuni)
rjb.test(zuni[, "Revenue"])
```

---

rlm.test   *Robust L1 Moment-Based (RLM) Goodness-of-Fit Test for the Laplace Distribution*

Description

Robust test for the Laplace distribution. Two options for calculating critical values, namely, approximated with Chi-square distribution and empirical, are available.

Usage

```
rlm.test(x, crit.values = c("chisq.approximation", "empirical"), N = 0)
```

Arguments

x	a numeric vector of data values.
crit.values	a character string specifying how the critical values should be obtained: approximated by the Chi-square distribution (default) or empirically.
N	number of Monte Carlo simulations for the empirical critical values.

Details

The test is based on a joint statistic using skewness and kurtosis coefficients. In particular, RLM uses the Average Absolute Deviation from the Median (MAAD), a robust estimate of standard deviation. See Gel (2010).

Value

A list of class `"htest"` with the following components:

statistic	the value of the test statistic.
parameter	the degrees of freedom.
p.value	the p-value of the test.
method	type of test was performed.
data.name	a character string giving the name of the data.

Author(s)

Kimihiro Noguchi, W. Wallace Hui, Yulia R. Gel

References

Gel YR (2010). "Test of fit for a Laplace distribution against heavier tailed alternatives." *Computational Statistics & Data Analysis*, **54**(4), 958–965. doi:10.1016/j.csda.2009.10.008.

See Also

sj.test, rjb.test, rqq, jarque.bera.test

Examples

```
## Laplace distributed data
x = rexp(100) - rexp(100)
rlm.test(x)
```

---

robust.mmm.test   *Robust Mudholkar–McDermott–Mudholkar Test for Ordered Variances*

Description

A test for a monotonic trend in variances (Mudholkar et al. 1995). The test statistic is based on a combination of the finite intersection approach and the two-sample t-test using Miller's transformation. By default, NAs are omitted.

Usage

```
robust.mmm.test(y, group, tail = c("right", "left", "both"))
```

Arguments

y	a numeric vector of data values.
group	factor of the data.
tail	the default option is "right", corresponding to an increasing trend in variances as the one-sided alternative; "left" corresponds to a decreasing trend in variances, and "both" corresponds to any (increasing or decreasing) monotonic trend in variances as the two-sided alternative.

Value

A list with the following elements:

T	the statistic and p-value of the test based on the Tippett p-value combination.
F	the statistic and p-value of the test based on the Fisher p-value combination.
N	the statistic and p-value of the test based on the Liptak p-value combination.
L	the statistic and p-value of the test based on the Mudholkar-George p-value combination.

Each of the list elements is a list of class "htest" with the following elements:

statistic	the value of the test statistic.
p.value	the p-value of the test.
method	type of test performed.
data.name	a character string giving the name of the data.

Author(s)

Kimihiro Noguchi, Yulia R. Gel

References

Mudholkar GS, McDermott MP, Mudholkar A (1995). "Robust finite-intersection tests for homogeneity of ordered variances." *Journal of Statistical Planning and Inference*, **43**(1-2), 185–195. doi:10.1016/03783758(94)00018Q.

See Also

neuhauser.hothorn.test, levene.test, lnested.test, ltrend.test, mma.test

Examples

```
data(pot)
robust.mmm.test(pot[, "obs"], pot[, "type"], tail = "left")$N
```

---

rqq   *Test of Normality Using RQQ Plots*

Description

Produce robust quantile-quantile (RQQ) and classical quantile-quantile (QQ) plots for graphical assessment of normality and optionally add a line, a QQ line, to the produced plot. The QQ line may be chosen to be a 45-degree line or to pass through the first and third quartiles of the data. NAs from the data are omitted.

Usage

```
rqq(
  y,
  plot.it = TRUE,
  square.it = TRUE,
  scale = c("MAD", "J", "classical"),
  location = c("median", "mean"),
  line.it = FALSE,
  line.type = c("45 degrees", "QQ"),
  col.line = 1,
  lwd = 1,
  outliers = FALSE,
  alpha = 0.05,
  ...
)
```

Arguments

y	the input data.
plot.it	logical. Should the result be plotted?
square.it	logical. Should the plot scales be square? The default is TRUE.
scale	the choice of a scale estimator, i.e., the classical or robust estimate of the standard deviation.
location	the choice of a location estimator, i.e., the mean or median.
line.it	logical. Should the line be plotted? No line is the default.
line.type	If line.it = TRUE, the choice of a line to be plotted, i.e., the 45-degree line or the line passing through the first and third quartiles of the data.
col.line	the color of the line (if plotted).
lwd	the line width (if plotted).
outliers	logical. Should the outliers be listed in the output?
alpha	significance level of outliers. If outliers = TRUE, then all observations that are less than the 100*alpha-th standard normal percentile or greater than the 100*(1-alpha)-th standard normal percentile will be listed in the output.
...	other parameters passed to the plot function.

Details

An RQQ plot is a modified QQ plot where data are robustly standardized by the median and robust measure of spread (rather than mean and classical standard deviation as in the basic QQ plots) and then are plotted against the expected standard normal order statistics (Gel et al. 2005; Weisberg 2005). Under normality, the plot of the standardized observations should follow the 45-degree line, or QQ line. Both the median and robust standard deviation are significantly less sensitive to outliers than mean and classical standard deviation and therefore are more preferable in many practical situations to assess graphically deviations from normality (if any). We choose median and MAD as a robust measure of location and spread for our RQQ plots since this standardization typically provides a clearer graphical diagnostics of normality. In particular, deviations from the QQ line are usually more noticeable in RQQ plots in the case of outliers and heavy tails. Users can also choose to plot the 45-degree line or the 1st-3rd quartile line (see the argument line.type). No line is the default.

Value

A list with the following numeric components:

x	the x-coordinates of the points that were/would be plotted.
y	the original data vector, i.e., the corresponding y-coordinates, including NAs (if any).

Author(s)

W. Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Gel Y, Miao W, Gastwirth JL (2005). "The importance of checking the assumptions underlying statistical analysis: graphical methods for assessing normality." *Jurimetrics*, **46**, 3.

Weisberg S (2005). *Applied Linear Regression*, 3 edition. John Wiley & Sons, Hoboken, NJ.

See Also

[rjb.test](), [sj.test](), [qqnorm](), [qqplot](), [qqline]()

Examples

```
## Simulate 100 observations from standard normal distribution:
y = rnorm(100)
rqq(y)

## Using Michigan data
data(michigan)
rqq(michigan)
```

---

runs.test   *Runs Test for Randomness*

Description

Performs the runs test for randomness (Mendenhall and Reinmuth 1982). Users can choose whether to plot the correlation graph or not, and whether to test against two-sided, negative, or positive correlation. NAs from the data are omitted.

Usage

```
runs.test(
  y,
  plot.it = FALSE,
  alternative = c("two.sided", "positive.correlated", "negative.correlated")
)
```

Arguments

y	a numeric vector of data values.
plot.it	logical. If TRUE, then the graph will be plotted. If FALSE (default), then it is not plotted.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "negative.correlated", or "positive.correlated".

Details

On the graph, observations that are less than the sample median are represented by red letters "A", and observations that are greater or equal to the sample median are represented by blue letters "B".

Value

A list of class `"htest"` with the following components:

`statistic`	the value of the standardized runs statistic.
`p.value`	the p-value for the test.
`data.name`	a character string giving the names of the data.
`alternative`	a character string describing the alternative hypothesis.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Mendenhall W, Reinmuth JE (1982). *Statistics for Management and Economics*, 4 edition. Duxbury, Boston, MA.

See Also

bartels.test

Examples

```
##Simulate 100 observations from an autoregressive model
## of the first order (AR(1))
y = arima.sim(n = 100, list(ar = c(0.5)))

##Test y for randomness
runs.test(y)
```

`sj.test`	*Test of Normality – SJ Test*

Description

Perform the robust directed test of normality, which is based on the ratio of the classical standard deviation S to the robust standard deviation J (Average Absolute Deviation from the Median, MAAD) of the sample data. See Gel et al. (2007).

Usage

```
sj.test(x, crit.values = c("t.approximation", "empirical"), N = 0)
```

Arguments

x	a numeric vector of data values.
crit.values	a character string specifying how the critical values should be obtained, i.e., approximated by the t-distribution (default) or empirically.
N	number of Monte Carlo simulations for the empirical critical values.

Value

A list of class `"htest"` with the following components:

statistic	the standardized test statistic.
p.value	the p-value.
parameter	the ratio of the classical standard deviation S to the robust standard deviation J.
data.name	a character string giving the name of the data.

Author(s)

Wallace Hui, Yulia R. Gel, Joseph L. Gastwirth, Weiwen Miao

References

Gel YR, Miao W, Gastwirth JL (2007). "Robust directed tests of normality against heavy-tailed alternatives." *Computational Statistics & Data Analysis*, **51**(5), 2734–2746. doi:10.1016/j.csda.2006.08.022.

See Also

rqq, rjb.test, jarque.bera.test

Examples

```
data(bias)
sj.test(bias)
```

---

symmetry.test	*Test of Symmetry*

Description

Perform test for symmetry about an unknown median. Users can choose among the Cabilio–Masaro test (Cabilio and Masaro 1996), the Mira test (Mira 1999), or the MGG test (Miao et al. 2006); and between using asymptotic distribution of the respective statistics or a distribution from m-out-of-n bootstrap (Lyubchich et al. 2016). Additionally to the general distribution asymmetry, the function allows to test for negative or positive skeweness (see the argument side). NAs from the data are omitted.

Usage

```
symmetry.test(
  x,
  option = c("MGG", "CM", "M"),
  side = c("both", "left", "right"),
  boot = TRUE,
  B = 1000,
  q = 8/9
)
```

Arguments

x	data to be tested for symmetry.
option	test statistic to be applied. The options include statistic by Miao et al. (2006) (default), Cabilio and Masaro (1996), and Mira (1999).
side	choice from the three possible alternative hypotheses: general distribution asymmetry (side = "both", default), left skewness (side = "left"), or right skewness (side = "right").
boot	logical value indicates whether m-out-of-n bootstrap will be used to obtain critical values (default), or asymptotic distribution of the chosen statistic.
B	number of bootstrap replications to perform (default is 1000).
q	scalar from 0 to 1 to define a set of possible m for the m-out-of-n bootstrap. Default q = 8/9. Possible m are then set as the values unique(round(n*(q^j)) greater than 4, where n = length(x) and j = c(0:20).

Details

If the bootstrap option is used (boot = TRUE), a bootstrap distribution is obtained for each candidate subsample size m. Then, a heuristic method (Bickel et al. 1997; Bickel and Sakov 2008) is used for the choice of optimal m. Specifically, we use the Wasserstein metric (Ruschendorf 2001) to calculate distances between different bootstrap distributions and select m, which corresponds to the minimal distance. See Lyubchich et al. (2016) for more details.

Value

A list of class `"htest"` with the following components:

`method`	name of the method.
`data.name`	name of the data.
`statistic`	value of the test statistic.
`p.value`	p-value of the test.
`alternative`	alternative hypothesis.
`estimate`	bootstrap optimal m (given in the output only if bootstrap was used, i.e., `boot = TRUE`).

Author(s)

Joseph L. Gastwirth, Yulia R. Gel, Wallace Hui, Vyacheslav Lyubchich, Weiwen Miao, Xingyu Wang (in alphabetical order)

References

Bickel PJ, Gotze F, van Zwet WR (1997). "Resampling fewer than n observations: gains, losses, and remedies for losses." *Statistica Sinica*, **7**, 1–31.

Bickel PJ, Sakov A (2008). "On the choice of m in the m out of n bootstrap and confidence bounds for extrema." *Statistica Sinica*, **18**(3), 967–985.

Cabilio P, Masaro J (1996). "A simple test of symmetry about an unknown median." *Canadian Journal of Statistics*, **24**(3), 349–361. doi:10.2307/3315744.

Lyubchich V, Wang X, Heyes A, Gel YR (2016). "A distribution-free m-out-of-n bootstrap approach to testing symmetry about an unknown median." *Computational Statistics & Data Analysis*, **104**, 1–9. doi:10.1016/j.csda.2016.05.004.

Miao W, Gel YR, Gastwirth JL (2006). "A new test of symmetry about an unknown median." In Hsiung A, Zhang C, Ying Z (eds.), *Random Walk, Sequential Analysis and Related Topics – A Festschrift in Honor of Yuan-Shih Chow*, 199–214. World Scientific Publisher, Singapore. doi:10.1142/9789812772558_0013.

Mira A (1999). "Distribution-free test for symmetry based on Bonferroni's measure." *Journal of Applied Statistics*, **26**(8), 959–972. doi:10.1080/02664769921963.

Ruschendorf L (2001). "Wasserstein metric." In Hazewinkel M (ed.), *Encyclopaedia of Mathematics*. Springer, Berlin.

Examples

```
data(zuni) #run ?zuni to see the data description
symmetry.test(zuni[,"Revenue"], boot = FALSE)
```

---

zuni	*The Zuni Data from the Law Case: Zuni Public School v. United States Department of Education*

Description

Number of students and available revenue per student in each school district in New Mexico.

Usage

```
data(zuni)
```

Format

A data frame with 89 observations on 3 variables: District, Revenue, and Mem (number of students).

Details

The Zuni data come from a law case "The Zuni Public School District No. 89, Gallup-McKinley County Public School District No. 1, Petitioners v. United States Department of Education" concerning whether the revenue per pupil satisfied the standard for "equal" expenditures per pupil in the state. This classification determines whether most of the federal money given to the state under the law goes to the state or to the local school districts.

Source

Gastwirth (2006).

References

Gastwirth JL (2006). "A 60 million dollar statistical issue arising in the interpretation and calculation of a measure of relative disparity: Zuni Public School District 89 v. US Department of Education." *Law, Probability and Risk*, **5**(1), 33–61. doi:10.1093/lpr/mgl019.

Index