Package 'goft'

October 13, 2022

Type Package

Title Tests of Fit for some Probability Distributions

Version 1.3.6

Date 2020-06-25

Author Elizabeth Gonzalez-Estrada, Jose A. Villasenor-Alva

Maintainer Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>

Description Goodness-of-fit tests for skew-normal, gamma, inverse Gaussian, log-normal, 'Weibull', 'Frechet', Gumbel, normal, multivariate normal, Cauchy, Laplace or double exponential, exponential and generalized Pareto distributions. Parameter estimators for gamma, inverse Gaussian and generalized Pareto distributions.

License GPL (>= 3)

Depends R (>= 2.10), stats, fitdistrplus, sn

LazyLoad yes

Repository CRAN

NeedsCompilation no

Date/Publication 2020-06-25 19:40:03 UTC

R topics documented:

---

cauchy_test   *Tests for the Cauchy distribution*

Description

Two tests for the Cauchy distribution hypothesis.

Usage

```
cauchy_test(x, N = 10^3, method = "transf")
```

Arguments

x	a numeric data vector containing a random sample of real numbers.
N	number of Monte Carlo samples used to approximate the p-value of the tests. Default is N = 10^3.
method	character string giving the name of the method to be used for testing the Cauchy distribution hypothesis. Two available options are `"transf"` and `"ratio"`.

Details

Option `"ratio"` performs a test for the Cauchy distribution based on the ratio of the maximum likelihood estimator for the scale parameter and the mean absolute deviation (Gonzalez-Estrada and Villasenor, 2018).

Option `"transf"` performs a test based on a data transformation to approximately exponentially distributed data (Villasenor and Gonzalez-Estrada, 2020).

Value

A list with class `"htest"` containing the following components.

statistic	the calculated value of the test statistic.
p.value	the approximated p-value of the test.
method	the character string "Test for the Cauchy distribution based on the ratio of two scale estimators".
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>

References

Gonzalez-Estrada, E. and Villasenor, J.A. (2018). An R package for testing goodness of fit: goft. *Journal of Statistical Computation and Simulation*, **88** 4, 726-751. https://doi.org/10.1080/00949655.2017.1404604

Villasenor, J.A. and Gonzalez-Estrada, E. (2020). Goodness of fit tests for Cauchy distributions using data transformations. In I. Ghosh, N. Balakrishnan and H.K.T. Ng. *Contributions of Barry C. Arnold to Statistical Science - Theory and Applications*. Springer.

Examples

```
x <- rnorm(20)      # simulating a data set from a normal distribution
cauchy_test(x)      # testing the Cauchy distribution hypothesis
```

| ev_test | *Tests for the extreme value distributions* |

Description

Correlation and ratio tests for Fisher-Tippet extreme value distributions.

Usage

```
ev_test(x, dist = "gumbel", method = "cor", N = 1000)
```

Arguments

x	a numeric data vector containing a random sample.
dist	the extreme value distribution to be tested. Options are "gumbel", "frechet" and "weibull" for testing the Gumbel, Frechet and Weibull distribution hypotheses. Default option is "gumbel".
method	the test to be used. Available options are "ratio" and "cor". Default is "cor".
N	number of Monte Carlo samples used to approximate the p-value of the test when "ratio" option is chosen. Default is N = 1000.

Details

Option "ratio" performs a test based on the ratio of two estimators for the variance of the Gumbel (type I extreme value) distribution (Gonzalez-Estrada and Villasenor, 2018).

Option "cor" performs a test based on the max-stability property of extreme value distributions (Gonzalez-Estrada and Villasenor, 2010). The sample size must lie between 20 and 250.

Value

A list with class `"htest"` containing the following components.

statistic   the calculated value of the test statistic.

p.value   the approximated p-value of the test.

method   a character string for the method used to test the null hypothesis.

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

Author(s)

Elizabeth Gonzalez-Estrada, Jose A. Villasenor

References

Gonzalez-Estrada, E. and Villasenor, J.A. (2018). An R package for testing goodness of fit: goft. *Journal of Statistical Computation and Simulation*, **88** 4, 726-751. `https://doi.org/10.1080/00949655.2017.1404604`

Gonzalez-Estrada, E. and Villasenor-Alva, J.A. (2010). A Goodness-of-Fit Test for Location-Scale Max-Stable Distributions. *Communications in Statistics: Simulation and Computation*, **39** 3, 557-562. `https://doi.org/10.1080/03610910903528293`

Examples

```
# Example 1:  testing the Gumbel distribution hypothesis using the correlation test
x <- -log(rexp(20))   # simulating a data set from the standard Gumbel distribution
ev_test(x)
```

---

exp_test   *Tests for exponentiality*

Description

Test based on a data trasformation and Cox-Oakes test for exponentiality.

Usage

```
exp_test(x, method = "transf", N = 10^3)
```

Arguments

x   a numeric data vector containing a random sample of positive real numbers.

method   test to be used. Tow available options are `"transf"` and `"ratio"`. Default option is `"transf"`.

N   number of Monte Carlo samples used to approximate p-values when `"transf"` option is chosen and the sample size is less than 200. Default is N = 10^3.

Details

Option "transf" performs a test based on a transformation to approximately uniformly distributed variables. If the sample size is larger than 200, the asymptotic null distribution of the test statistic is used to approximate the p-value; otherwise, it is approximated by Monte Carlo simulation (Villasenor and Gonzalez-Estrada, 2020).

Option "ratio" performs Cox and Oakes (1984) test.

Value

A list with class "htest" containing the following components.

statistic	the calculated value of the test statistic.
p.value	approximated p-value of the test.
method	a character string giving the name of the method used for testing exponentiality.
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada, Jose A. Villasenor

References

Villasenor, J.A. and Gonzalez-Estrada, E. 2020. On testing exponentiality based on a new estimator for the scale parameter. *Brazilian Journal of Probability and Statistics* (accepted for publication).

Cox, D.R., Oakes, D. 1984. Analysis of Survival Data. Chapman and Hall/CRC.

Examples

```
# Testing exponentiality on a simulated random sample from the exponential distribution
x <- rexp(20)
exp_test(x)
```

---

gamma_fit	*Fitting the Gamma distribution to data*

Description

Fits a Gamma distribution to a random sample of positive real numbers using Villasenor and Gonzalez-Estrada (2015) parameter estimators.

Usage

```
gamma_fit(x)
```

Arguments

x   a numeric data vector containing a random sample of positive real numbers.

Details

The Gamma distribution with shape and scale parameters is considered. The scale parameter is estimated by the unbiased sample estimator of the covariance of X and log(X). The shape parameter is estimated by the ratio of the sample mean of X and the scale parameter estimator.

Value

Shape and scale parameter estimates.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor, J.A. and Gonzalez-Estrada, E. (2015). A variance ratio test of fit for Gamma distributions. *Statistics and Probability Letters*, **96** 1, 281-286. http://dx.doi.org/10.1016/j.spl.2014.10.001

See Also

[gamma_test](#) for testing the Gamma distribution hypothesis.

Examples

```
# Fitting a gamma distribution to the logarithm of variable Loss contained in
# the danishuni data set
library(fitdistrplus)
data(danishuni)
logLoss <- log(danishuni$Loss)   # logarithm of Loss variable
logLoss <- logLoss[logLoss > 0]  # observations > 0
gamma_fit(logLoss)
```

---

gamma_test   *Test for the Gamma distribution*

Description

Test of fit for the Gamma distribution with unknown shape and scale parameters based on the ratio of two variance estimators (Villasenor and Gonzalez-Estrada, 2015).

Usage

```
gamma_test(x)
```

Arguments

x   a numeric data vector containing a random sample of positive real numbers.

Details

The test statistic is the ratio of two variance estimators, namely, the sample variance and the moments estimator obtained by Villasenor and Gonzalez-Estrada (2015), which is the product of the sample mean of X and the sample covariance of X and log(X).

The asymptotic null distribution of the test statistic is used to approximate p-values.

NOTE: the unbiased sample covariance estimator is used to compute the test statistic.

Value

A list with class `"htest"` containing the following components.

statistic	the calculated value of the test statistic.
p.value	the approximated p-value of the test.
method	the character string "Test of fit for the Gamma distribution".
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor, J.A. and Gonzalez-Estrada, E. (2015). A variance ratio test of fit for Gamma distributions. *Statistics and Probability Letters*, **96** 1, 281-286. http://dx.doi.org/10.1016/j.spl.2014.10.001

See Also

gamma_fit for fitting a Gamma distribution to data.

Examples

```
# Testing the gamma distribution hypothesis on the logarithm of variable Loss
# of the danishuni data set
library(fitdistrplus)
data(danishuni)
logLoss <- log(danishuni$Loss)   # logarithm of Loss variable
logLoss <- logLoss[logLoss > 0] # observations > 0
gamma_test(logLoss)
```

---

goats   *Zoometric measurements of goats*

Description

Zoometric measurements of 27 week old creole goats collected by Dorantes-Coronado (2013).

Usage

```
data(goats)
```

Format

A data frame with 52 rows and 7 columns containing measurements (in kilograms and centimeters) on the following variables.

```
body.weight
```
```
body.length
```
```
trunk.length
```
```
withers.height
```
```
thoracic.perimeter
```
```
hip.length
```
```
ear.length
```

Source

Dorantes-Coronado (2013).

References

Dorantes-Coronado, E.J. (2013). Estudio preliminar para el establecimiento de un programa de mejoramiento genetico de cabras en el Estado de Mexico. Ph.D. Thesis. Colegio de Postgraduados, Mexico.

Examples

```
data(goats)
plot(goats)
```

---

gp_fit   *Fitting the generalized Pareto distribution to data*

Description

Fits a generalized Pareto distribution (gPd) to a random sample using either the asymptotic maximum likelihood method (amle) or the combined estimation method (Villasenor-Alva and Gonzalez-Estrada, 2009).

Usage

```
gp_fit(x, method)
```

Arguments

x   a numeric data vector containing a random sample of positive real numbers.

method   a character string giving the name of the parameter estimation method to be used. There are two available methods: "combined" and "amle". Use "combined" for fitting a gPd with negative shape parameter. Use "amle" for fitting a gPd with non-negative shape parameter.

Details

The distribution function of the gPd is provided in the details section of the [gp_test](#) function.

Value

Parameter estimates.

Author(s)

Elizabeth Gonzalez-Estrada, Jose A. Villasenor-Alva

References

Villasenor-Alva, J.A. and Gonzalez-Estrada, E. (2009). A bootstrap goodness of fit test for the generalized Pareto distribution. *Computational Statistics and Data Analysis*,**53**,11,3835-3841. http://dx.doi.org/10.1016/j.csda.2009.04.001

See Also

[gp_test](#) for testing the gPd hypothesis.

Examples

```
# Example 1: fitting a gPd to the excesses above the threshold 0.165 ppm of the ozone
# levels provided in the o3 data set.
data(o3)
o3levels <- o3$ozone_level - 0.165      # ozone levels minus the threshold 0.165 ppm
gp_fit(o3levels, method = "combined")   # fitting the gPd using the "combined" method

# Example 2: fitting a gPd to a simulated data set from the uniform distribution,
# which is a gPd with shape parameter = -1 and scale parameter = 1.
x <-  runif(100)       # simulating a uniform random sample
gp_fit(x, "combined")  # fitting a gPd to  x  using the "combined" method
```

---

gp_test   *Bootstrap test for the generalized Pareto distribution*

Description

Test of fit for the generalized Pareto distribution (gPd) with unknown parameters by Villasenor-Alva and Gonzalez-Estrada (2009).

Usage

```
gp_test(x, B = 999)
```

Arguments

x   numeric data vector containing a random sample of positive real numbers.

B   number of bootstrap samples used to approximate p-values. Default is B=999.

Details

This bootstrap test for the null hypothesis H_0 : a random sample has a gPd with unknown shape parameter γ is an intersection-union test for the hypotheses H_0^- : a random sample has a gPd with $\gamma < 0$, and H_0^+ : a random sample has a gPd with $\gamma >= 0$. Thus, heavy and non-heavy tailed gPd's are included in the null hypothesis. The parametric bootstrap is performed on γ for each of the two hypotheses.

The gPd function with unknown shape and scale parameters γ and σ is given by

$$F(x) = 1 - \left[1 + \frac{\gamma x}{\sigma}\right]^{-1/\gamma},$$

where γ is a real number, $\sigma > 0$ and $1 + \gamma x/\sigma > 0$. When $\gamma = 0$, F(x) becomes the exponential distribution with scale parameter σ:

$$F(x) = 1 - exp\left(-x/\sigma\right).$$

Value

A list with class `"htest"` containing the following components.

p.value	an approximated p-value of the test using parametric bootstrap.
method	the character string "Bootstrap test of fit for the generalized Pareto distribution".
data.name	a character string giving the name of the data set.
pvalues	approximated p-values of the tests for H_0^- and H_0^+

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor-Alva, J.A. and Gonzalez-Estrada, E. (2009). A bootstrap goodness of fit test for the generalized Pareto distribution. *Computational Statistics and Data Analysis*,**53**,11,3835-3841. http://dx.doi.org/10.1016/j.csda.2009.04.001

See Also

[gp_fit](gp_fit) for fitting a gPd to data.

Examples

```
# Testing the gPd hypothesis on the excesses above the threshold 0.165 ppm of the ozone
# levels given in the o3 data set
data(o3)
o3levels <- o3$ozone_level - 0.165  # ozone levels minus the threshold 0.165 ppm
gp_test(o3levels)                   # testing the gPd hypothesis
```

---

ig_fit	*Fitting the Inverse Gaussian distribution to data*

Description

Maximum likelihood estimators for the Inverse Gaussian distribution based on a random sample.

Usage

```
ig_fit(x)
```

Arguments

x	a numeric data vector containing a random sample of positive real numbers.

Value

Parameter estimates.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

See Also

[ig_test](#) for testing the Inverse Gaussian distribution hypothesis.

Examples

```
x <-  rgamma(50, 10)    # simulating a random sample from the gamma distribution
ig_fit(x)               # fitting an inverse Gaussian distribution to x
```

---

ig_test   *Tests for the Inverse Gaussian distribution*

Description

Implementation of three tests of fit for Inverse Gaussian distributions with unknown parameters (see Details section).

Usage

```
ig_test(x, method = "transf")
```

Arguments

x   a numeric data vector containing a random sample of positive real numbers.

method   character string giving the name of the method to be used for testing the Inverse Gaussian hypothesis. Two available options are "transf" and "ratio".

Details

Option "transf" performs two tests based on data transformations. One test transforms the observations to approximately normally distributed observations and then uses Shapiro-Wilk test for asssessing univariate normality (Villasenor et al., 2019). The other test is based on a transformation of data to gamma variables with shape parameter equal to 1/2 and uses Anderson-Darling test for testing the gamma distribution (Villasenor and Gonzalez-Estrada, 2015). The right-hand side tail of the null distribution for Anderson-Darling test statistic is approximated by an inverse Gaussian distribution in order to get approximated p-values.

Option "ratio" performs a test based on the ratio of two variance estimators. It is based on the asymptotic distribution of the test statistic. The use of this test is appropriate for large sample sizes (Villasenor and Gonzalez-Estrada, 2015).

Value

A list with class `"htest"` containing the following components.

`statistic`	the calculated value of the test statistic.
`p.value`	the approximated p-value of the test.
`method`	a character string giving the name of the method used for testing the null hypothesis.
`data.name`	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor, J.A. and Gonzalez-Estrada, E. (2015). Tests of fit for Inverse Gaussian distributions. *Statistics and Probability Letters*, **105**, 189-194. http://dx.doi.org/10.1016/j.spl.2015.06.017

Villasenor, J.A. Gonzalez-Estrada, E. and Ochoa, A. (2019). On Testing the Inverse Gaussian Distribution Hypothesis. *Sankhya B: The Indian Journal of Statistics*. **81**, 60-74. https://doi.org/10.1007/s13571-017-0148-8

Examples

```
# Example 1: testing the inverse Gaussian distribution hypothesis  using "transf"
# option on the compressive strength variable of the strength data set.
data("strength")
y <- strength$cstrength        # compressive strength
ig_test(y, method = "transf")

# Getting the p-value of the test based on a transformation to gamma variables
ig_test(y)[[2]]$p.value

# Example 2: testing the inverse Gaussian distribution hypothesis  using "ratio"
# option on a simulated data set.
x <- rlnorm(500)    # simulating a random sample from the lognormal distribution
ig_test(x, method = "ratio")
```

---

laplace_test	*Tests for the Laplace or double exponential distribution*

Description

Transformation and ratio tests for the Laplace distribution by Gonzalez-Estrada and Villasenor (2016).

Usage

```
laplace_test(x, method = "transf", N = 10^5)
```

Arguments

x	a numeric data vector containing a random sample of real numbers.
method	the type of test to be performed. Two available options are "transf" and "ratio". Default option is "transf".
N	number of Monte Carlo samples used to approximate the p-value of the test when the "ratio" option is chosen and the sample size is less than 500. Default is N = 10^5.

Details

When "transf" option is chosen, a transformation to approximately exponential random variables is performed and the exponentiality hypothesis is assessed using Anderson-Darling test.

When "ratio" option is chosen, a test based on the ratio of two estimators of the scale parameter is performed. For samples of size $n < 500$, the p-value of this test is approximated by Monte Carlo simulation. Otherwise, it is approximated by the standard normal cumulative distribution function.

Value

A list with class "htest" containing the following components.

statistic	the calculated value of the test statistic.
p.value	approximated p-value of the test.
method	a character string giving the name of the method used for testing the null hypothesis.
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor

References

Gonzalez-Estrada, E. and Villasenor, J.A. (2016). A ratio goodness-of-fit test for the Laplace distribution. *Statistics and Probability Letters*, **119**, 30-35. https://doi.org/10.1016/j.spl.2016.07.003

Examples

```
# Example 1:  testing the Laplace distribution hypothesis using "transf" option
x <- rnorm(50)   # simulating a random sample from a normal distribution
laplace_test(x)

# Example 2: testing the Laplace distribution hypothesis using "ratio" option
x <- rt(60,4)    # simulating a random sample from Student's t distribution with 4 d.f.
```

```
laplace_test(x, method = "ratio")
```

---

lnorm_test   *Test for the lognormal distribution*

Description

Test for the lognormal distribution based on a data transformation to normal observations.

Usage

```
lnorm_test(x)
```

Arguments

x   a numeric data vector containing a random sample of positive observations.

Details

Shapiro-Wilk test is used for testing normality on the transformed observations.

Value

A list with class `"htest"` containing the following components.

statistic   the calculated value of the test statistic.

p.value   an approximated p-value of the test.

method   the character string "Test for the lognormal distribution based on a transformation to normality".

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

Author(s)

Elizabeth Gonzalez-Estrada, Jose A. Villasenor

See Also

Other tests for some alternative distributions to the lognormal model are impremented in functions
`ig_test`, `gamma_test` and `weibull_test`.

Examples

```
# Testing the lognormal distribution hypothesis  on the compressive strength variable
# of the strength data set.
data("strength")
x  <- strength$cstrength   # compressive strength
lnorm_test(x)              # testing the lognormal distribution hypothesis
```

---

mvshapiro_test   *Shapiro-Wilk test for multivariate normality*

Description

A generalization of Shapiro-Wilk test for multivariate normality (Villasenor-Alva and Gonzalez-Estrada, 2009).

Usage

```
mvshapiro_test(X)
```

Arguments

X   a numeric data matrix with d columns and n rows.

Details

Sample size (n) must be larger than vector dimension (d).

When d = 1, `mvshapiro_test(X)` produces the same results as `shapiro.test(X)`.

Value

A list with class `"htest"` containing the following components.

statistic   the value of the generalized Shapiro-Wilk statistic for testing multivariate normality.

p.value   an approximated p-value of the test.

method   the character string "Generalized Shapiro-Wilk test for multivariate normality".

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

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor-Alva, J.A. and Gonzalez-Estrada, E. (2009). A generalization of Shapiro-Wilk's test for multivariate normality. *Communications in Statistics: Theory and Methods*, **38** 11, 1870-1883. http://dx.doi.org/10.1080/03610920802474465

See Also

shapiro.test and normal_test for testing univariate normality.

Examples

```
# Example 1:  Testing multivariate normality on iris.virginica

# iris.virginica contains a set of measurements corresponding to
# Iris virginica of famous  iris data set.

iris.virginica <- as.matrix(iris[iris$Species == "virginica", 1:4], ncol = 4)
mvshapiro_test(iris.virginica)


# Example 2:  Testing multivariate normality on the goats dataset
data(goats)
mvshapiro_test(as.matrix(goats))
```

---

normal_test   *Test for normality*

Description

Correlation test of fit for normality based on the Levy characterization (Villasenor-Alva and Gonzalez-Estrada, 2015).

Usage

```
normal_test(x, method = "cor")
```

Arguments

x	a numeric data vector containing a random sample of size n.
method	a character string giving the name of the test to be used. So far the only option is "cor".

Details

Sample size (n) must be an integer lying between 10 and 400.

The Levy characterization of the normal distribution states that sums of independent normal random variables are also normal random variables. Based on this property, the normality assumption might be judged by comparing sums of pairs of observations from a random sample to quantiles of the standard normal distribution. If normality holds then the pairs of sums and quantiles should lie on a straight line approximately. A formal test for normality is obtained when such a comparison is based on the sample correlation coefficient of sums and quantiles.

Value

A list with class "htest" containing the following components.

statistic   the calculated value of the test statistic.

p.value	an approximated p-value of the test.
method	the character string "Correlation test for normality".
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada <egonzalez@colpos.mx>, Jose A. Villasenor-Alva

References

Villasenor-Alva, J.A. and Gonzalez-Estrada, E. (2015). A correlation test for normality based on the Levy characterization. *Communications in Statistics: Simulation and Computation*, **44** 5, 1225-1238. http://dx.doi.org/10.1080/03610918.2013.810261

See Also

shapiro.test

Examples

```
data(goats)                  # loading the "goats" data set
apply(goats,2,normal_test)   # testing normality on each variable of the "goats" data set
```

---

o3   *Ozone concentrations over 0.165 ppm*

Description

Ozone levels in parts per million (ppm) registered in the southwest of Mexico City from January 1, 2008 to April 30th, 2016, which exceed the 0.165 ppm threshold.

Usage

```
data(o3)
```

Format

A data frame with 47 observations on the following 2 variables.

date date in the format yyyy-mm-dd
ozone_level a numeric vector containing the ozone levels above 0.165 ppm

Source

Atmospheric Monitoring Office of Mexico City.

Examples

```
data(o3)
plot(o3)
```

---

o3max	*Monthly maximum ozone concentrations*

Description

Maximum ozone concentrations (in ppm) per month registered in the southwest of Mexico City from 2008 to 2015 during the months of March to May.

Usage

```
data(o3max)
```

Format

A data frame with 24 observations on the following 2 variables.

month months in which observations were registered.

o3max_level a numeric vector containing the maximum ozone leveles per month.

Source

Atmospheric Monitoring Office of Mexico City.

Examples

```
data(o3max)
plot(o3max)
```

---

sn_test	*Test for skew normality*

Description

Transformation test of fit for skew-normality (Gonzalez-Estrada and Cosmes, 2019).

Usage

```
sn_test(x, method = "transf")
```

Arguments

x	a numeric data vector containing a random sample of size n.
method	a character string giving the name of the test to be used. So far the only option is "transf".

Details

The random sample is transformed to approximately normal variables. Shapiro-Wilk test is used for testing normality.

The sample size should be larger than 50 in order to have a reasonable control of the type I error probability.

Value

A list with class `"htest"` containing the following components.

`p.value`	an approximated p-value of the test.
`method`	the character string "Shapiro-Wilk test for skew-normal distributions".
`data.name`	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada (<egonzalez@colpos.mx>)

References

Gonzalez-Estrada, E. and Cosmes, W. (2020). Shapiro-Wilk test for skew normal distributions based on data transformations. *Journal of Statistical Computation and Simulation*, **89** 17, 3258-3272. https://doi.org/10.1080/00949655.2019.1658763

Examples

```
data(strength)            # loading the "strength" data set
y  <- strength$strain
sn_test(y)                # testing skew normality on the strain variable
```

---

strength	*Compressive strength of maize seeds*

Description

Compressive strength and strain of maize seeds.

Usage

```
data("strength")
```

Format

A data frame with 90 observations on the following 2 variables.

strain a numeric vector giving the relative change in length under compression stress in millimeters.

cstrength a numeric vector giving the compressive strength in Newtons.

Details

These data correspond to maize seeds with floury endosperm and 8% of moisture.

Source

Mancera-Rico, A. (2014).

References

Mancera-Rico, A. (2014). Contenido de humedad y tipo de endospermo en la resistencia a compresion en semillas de maiz. Ph.D. Thesis. Colegio de Postgraduados, Mexico.

Examples

```
data(strength)
plot(strength)      # plot of "strain" versus "cstrength"
```

---

weibull_test	*Test for the Weibull distribution*

Description

Test for the Weibull distribution based on a data transformation to observations with type I extreme value (Gumbel) distribution.

Usage

```
weibull_test(x, method = "transf", N = 1000)
```

Arguments

x	numeric data vector containing a random sample of positive observations.
method	test to be used. So far, `"transf"` is the only option and it comes by default.
N	number of Monte Carlo samples used to approximate the p-value of the test. Default is N = 1000.

Details

The `"ratio"` test provided by function [ev_test](#) for the type I extreme value distribution is applied to the transformed observations.

Value

A list with class `"htest"` containing the following components.

statistic	the calculated value of the test statistic.
p.value	an approximated p-value of the test.
method	the character string "Test for the Weibull distribution".
data.name	a character string giving the name of the data set.

Author(s)

Elizabeth Gonzalez-Estrada, Jose A. Villasenor

See Also

Other tests for positive datasets are provided by functions `ig_test`, `gamma_test` and `lnorm_test`.

Examples

```
# Testing the Weibull distribution hypothesis on the maximum ozone levels given
# in the o3max data set.
data(o3max)
x <- o3max[ ,2]     # maximum ozone levels
weibull_test(x, N = 10000)
```

Index