Package 'goft'

July 2, 2014

Type Package

Title Tests of fit for some probability distributions

Version 1.0

Date 2013-10-18

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

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

Description This package implements some tests of fit for the normal, Gumbel (type I extreme value distribution), multivariate normal and generalized Pareto distributions.

License GPL (>= 3)

Depends stats, gPdtest, mvShapiroTest

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2013-12-04 07:29:16

R topics documented:

`gpd.test` *Bootstrap goodness-of-fit test for the generalized Pareto distribution*

Description

This function computes the bootstrap goodness-of-fit test by Villasenor-Alva and Gonzalez-Estrada (2009) for testing the null hypothesis H_0 : a random sample has a generalized Pareto distribution (gPd) with unknown shape parameter γ, which is a real number.

Usage

```
gpd.test(x,J)
```

Arguments

`x`	numeric data vector containing a random sample from a distribution function with support on the positive real numbers.
`J`	number of bootstrap samples. This is an optional argument. Default `J=999`.

Details

The bootstrap goodness-of-fit test for the gPd is an intersection-union test for testing 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.

We consider the distribution function of the gPd with unknown shape and scale parameters γ and σ 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 $\gamma = 0$, we have the exponential distribution with scale parameter σ:

$$F(x) = 1 - exp\,(-x/\sigma)$$

Value

A list with the following components.

`boot.test`	a list with class `"htest"` containing the p-value of the test, the name of the data set, and the character string "Bootstrap goodness-of-fit test for the generalized Pareto distribution".
`p.values`	the p-values of the tests of the hypotheses H_0^- and H_0^+ described above.

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.

See Also

`gpd.fit` for fitting a gPd to data, `rgp` for generating gPd random numbers.

Examples

```
x <-  rgp(20,shape = 1)    ## generating a random sample of size 20 from a gPd
gpd.test(x)                 ## testing the gPd hypothesis on x
```

---

gumbel.test *Correlation test for the Gumbel distribution*

Description

This function implements the correlation test of fit for the Gumbel distribution (also known as type I extreme value distribution) based on the max-stability property (Gonzalez-Estrada and Villasenor-Alva, 2010).

Usage

```
gumbel.test(x)
```

Arguments

`x`	Numeric data vector containing a random sample of size n.

Details

Sample size (n) must be an integer between 20 and 250.

The max-stability property of the Gumbel distribution states that the maximum of a number of Gumbel random variables is also a Gumbel random variable.

The test statistic is the sample correlation of the maximum of pairs of observations and quantiles of the Gumbel distribution.

Value

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

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

References

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.

Examples

```
x <- -log(rexp(20))   # generating a pseudo random sample of size 20 from
                      # the standard Gumbel distribution
gumbel.test(x)        # testing the Gumbel hypothesis on x
```

`mvShapiro.Test` *Shapiro-Wilk test for multivariate normality*

Description

Given a d-dimensional random sample of size n, this function computes the test statistic and p-value of the generalized Shapiro-Wilk test for multivariate normality proposed by Villasenor-Alva and Gonzalez-Estrada (2009).

Usage

```
mvShapiro.Test(X)
```

Arguments

`X`	Numeric data matrix with d columns (vector dimension) and n rows (sample size).

Details

n must be larger than 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`	the 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.

See Also

shapiro.test

Examples

```
X <-  matrix(rnorm(40),ncol=2)    # Generating a two dimensional random sample of size 20
mvShapiro.Test(X)                 # Testing multivariate normality on X

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

iris.virginica <- as.matrix(iris[iris$Species == "virginica",1:4],ncol=4)
mvShapiro.Test(iris.virginica)    # Testing multivariate normality on iris.virginica
```

---

normal.test *Correlation test for univariate normality*

Description

This function implements the correlation test for normality based on the Levy characterization by Villasenor-Alva and Gonzalez-Estrada (2013).

Usage

```
normal.test(x)
```

Arguments

x	Numeric data vector containing a random sample of size n.

Details

Sample size (n) must be an integer 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 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 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	the 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.

References

Villasenor-Alva, J.A. and Gonzalez-Estrada, E. (2013). A correlation test for normality based on the Levy characterization. *Communications in Statistics: Simulation and Computation* (to appear).

See Also

shapiro.test

Examples

```
x <- rnorm(20)        # generating a pseudo random sample of size 20 from
                      # the standard normal distribution
normal.test(x)        # testing normality on x
```

Index