Package 'GOFSN'

January 2, 2012

Type Package

Title Goodness-of-fit tests for the family of skew-normal models

Version 1.0

Date 2009-04-24

Author Veronica Paton Romero

Maintainer Veronica Paton Romero <v.paton@alumnos.urjc.es>

Depends R(>= 2.7.0), sn

Description GOFSN is a package that implements a method for checking
if a skew-normal model fits the observed dataset, when all
parameters are unknown. While location and scale parameters are
estimated by moment estimators, the shape parameter is integrated with respect to the prior predictive distribution,as proposed in (BOX, 1980). A default and proper prior on
skewness parameter is used to obtain the prior predictive
distribution, as proposed in (CABRAS, CASTELLANOS, 2008).
Goodness-of-fit tests, here proposed, depend only on sample
size and exhibit full agreement between nominal and actual
size. This package implements EDF statistics
Kolmogorov-Smirnov(D), Cram\{}'er-von Mises(W2) and proposes some
simple algorithms (SimulD,SimulW2) to approximate their
respective marginal predictive distributions. It also has
functions (ks.sn,W2.sn) that calculate the p-value on observed data.

License GPL (>= 2)

LazyLoad yes

Repository CRAN

Date/Publication 2009-06-22 19:00:04

R topics documented:

`GOFSN-package` *Package GOFSN: summary information*

Description

This package proposes a series of goodness-of-fit tests for the family of skew-normal models when all parameters are unknown. It provides functions that analyse if data are compatible with the assumed model.

Details

Package:	GOFSN
Type:	Package
Version:	1.0
Date:	2009-04-24
License:	GPL (version 2 or later)
LazyLoad:	yes

Functions

The package includes functions to analyze the fit of a univariate distribution to a proposed data set: `D`, `W2`, `SimulD`, `SimulW2`, `ks.sn`, `W2.sn`.

It also includes some objects necessary for the above functions: `prior.lambda`, `D.quantiles`, `W2.quantiles`.

Requeriments

R(>=2.7.0), sn

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

Examples

```
x<-c(2,5,10,4,3,8)
D.data<-D(x)
W2.data<-W2(x)
```

D *The Kolmogorov-Smirnov statistic D*

Description

This statistic measures the difference between the empirical distribution, Fn, and a theoretical one, F, which must be some element of the SN class. It is only valid for continuos variables.

Usage

```
D(data)
```

Arguments

data	a numeric vector with observed data

Details

Central parametrization is used, and the unknown parameters are estimated by: MLE of skewness, sample mean for mu, and sample standard desviation for sigma.

Value

x	a numeric value with the disparity between Fn and F

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

W2, prior.lambda

Examples

```
data(ais, package="sn")
attach(ais)
data=bmi
D.data=D(data)
```

`D.quantiles` *Matrix D for all sizes and quantiles*

Description

This matrix provides critical quantiles of the prior predictive distribution `m(t)` for the EDF stadistic D and different sizes `n`.

Usage

```
data(D.quantiles)
```

Format

A data frame with 51 rows and 6 columns. Rows represent the sample size (5,6,...,49,50,60,70,80,90,100) and columns represent the critical quantiles (0.10,0.05,0.025,0.010,0.005,0.001).

Details

Quantiles of the prior predictive distribution for statistic D are approximated by Monte Carlo simulations, using 100.000 draws from this predictive distribution. Details of calculation are given in Section 3.2 from article Cabras,Castellanos (2008).

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

`SimulD`, `ks.sn`

Examples

```
data(D.quantiles)
D.size5=D.quantiles[1,]
```

ks.sn *Observed Kolmogorov-Smirnov statistic D and aproximate prior predictive p-value*

Description

This function calculates the Kolmogorov-Smirnov statistic in an observed data set and approximate the prior predictive p-value comparing the statistic with the approximated quantiles contained in `D.quantiles`.

Usage

```
ks.sn(data)
```

Arguments

`data`	a numeric vector with the observed data

Details

First, this function calculates the statistic D on the observed data, after this value is compared with the appropiate row of the matrix `D.quantiles` that provides critical quantiles of the prior predictive distribution `m(t)` for the EDF stadistic D and different sample sizes.

Value

`x`	a numeric value with the observed stadistic D and the p-value associated.

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

`W2.sn`

Examples

```
data(ais, package="sn")
attach(ais)
data=bmi
data(prior.lambda)
data(D.quantiles)
ks.sn(data)
```

`lambda.to.gamma` *Converts lambda values to gamma values*

Description

For a given value of lambda, parameter of skewness in direct parametrization, the function finds the corresponding value of gamma. This function is necessary for other functions as `SimulD` and `SimulW2`.

Usage

```
lambda.to.gamma(lambda)
```

Arguments

`lambda`	A numeric vector of simulated values of lambda.

Value

`gamma`	A numeric vector of the corresponding gamma values.

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

`SimulD`, `SimulW2`

Examples

```
data(prior.lambda)
lambdas=sample(prior.lambda,10000)
gammas=lambda.to.gamma(lambdas)
```

prior.lambda *Vector of simulations from the prior distribution for lambda*

Description

This vector contains values simulated from the Jeffrey's prior for lambda.

Usage

```
data(prior.lambda)
```

Format

A vector of length 1000000

Details

Jeffrey's prior has been simulated using a numerical approximation of the prior density and through the inverse method.

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

Examples

```
data(prior.lambda)
lambdas=sample(prior.lambda, 10000)
```

SimulD *Algorithm to derive the marginal prior predictive distribution of the EDF stadistic D*

Description

This function approximates the predictive prior distribution of statistic D for the SN model. This is achieved by integrating out gamma with respect to the Jeffreys' prior distribution.

Usage

```
SimulD(n, nrep)
```

Arguments

n	size
nrep	number of draws from the prior predictive distribution for gamma used to approximate h(t)

Details

h(t) is the marginal prior predictive distribution of D obtained by integrating gamma with respect to the Jeffreys'prior. To approximate h(t) for a particular n we used M=1000000 draws from the prior using the following 3 steps algorithm:

Step 1 Draw gamma(1,..,m,...M) approx pi(gamma), for M=1000000;
Step 2 for each gamma(m) generate a random sample of size n from the SN model with mu=0 sigma=1 and gamma=gamma(m);
Step 3 for the mth sample calculate the stadistic D on the random sample.

Although the Kolmogorov-Smirnov test in ks.sn makes use of the approximated quantiles for a set of sample sizes, with this function it is possible to approximate h(t) and to obtain its quantiles for any sample size n.

Value

y	a vector of simulations that contains a sample of values that aproximate h(t) for a particular n

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

SimulW2,prior.lambda

Examples

```
data(prior.lambda)
simulationsD=SimulD(5,10)
```

SimulW2 *Algorithm to derive the marginal prior predictive distribution of the EDF stadistic W2*

Description

This function approximates the predictive prior distribution of statistic D for the SN model. This is achieved by integrating out gamma with respect to the Jeffreys' prior distribution.

Usage

```
SimulW2(n, nrep)
```

Arguments

n	size
nrep	number of draws from the prior predictive distribution for gamma used to approximate h(t)

Details

h(t) is the marginal prior predictive distribution of D obtained by integrating gamma with respect to the Jeffreys'prior. To approximate h(t) for a particular n we used M=1000000 draws from the prior using the following 3 steps algorithm:

Step 1 Draw gamma(1,..,m,...M) approx pi(gamma), for M=1000000;
Step 2 for each gamma(m) generate a random sample of size n from the SN model with mu=0 sigma=1 and gamma=gamma(m);
Step 3 for the mth sample calculate the stadistic W2 on the random sample.

Although the Kolmogorov-Smirnov test in W2.sn makes use of the approximated quantiles for a set of sample sizes, with this function it is possible to approximate h(t) and to obtain its quantiles for any sample size n.

Value

y	a vector of simulations that contains a sample of values that aproximate h(t) for a particular n

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

SimulD,prior.lambda

Examples

```
data(prior.lambda)
simulationsW2=SimulW2(5,10)
```

W2 *The Kolmogorov-Smirnov statistic W2*

Description

This statistic measures the difference between the empirical distribution, Fn, and a theoretical one, F, which must be some element of the SN class. It is only valid for continuos variables.

Usage

```
W2(data)
```

Arguments

data	a numeric vector with observed data

Details

Central parametrization is used, and the unknown parameters are estimated by: MLE of skewness, sample mean for mu, and sample standard desviation for sigma.

Value

x	a numeric value with the disparity between Fn and F

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

W2, prior.lambda

Examples

```
data(ais, package="sn")
attach(ais)
data=bmi
W2.data=W2(data)
```

W2.quantiles *Matrix W2 for all sizes and quantiles*

Description

This matrix provides critical quantiles of the prior predictive distribution `m(t)` for the EDF stadistic W2 and different sizes `n`.

Usage

```
data(W2.quantiles)
```

Format

A data frame with 51 rows and 6 columns. Rows represent the sample size (5,6,...,49,50,60,70,80,90,100) and columns represent the critical quantiles (0.10,0.05,0.025,0.010,0.005,0.001).

Details

Quantiles of the prior predictive distribution for statistic W2 are approximated by Monte Carlo simulations, using 100.000 draws from this predictive distribution. Details of calculation are given in Section 3.2 from article Cabras,Castellanos (2008).

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

`SimulW2`, `W2.sn`

Examples

```
data(W2.quantiles)
W2.size5=W2.quantiles[1,]
```

W2.sn *Observed Kolmogorov-Smirnov statistic W2 and aproximate prior predictive p-value*

Description

This function calculates the Kolmogorov-Smirnov statistic in an observed data set and approximate the prior predictive p-value comparing the statistic with the approximated quantiles contained in `W2.quantiles`.

Usage

```
W2.sn(data)
```

Arguments

data	a numeric vector with the observed data

Details

First, this function calculates the statistic W2 on the observed data, after this value is compared with the appropiate row of the matrix `W2.quantiles` that provides critical quantiles of the prior predictive distribution `m(t)` for the EDF stadistic W2 and different sample sizes.

Value

x	a numeric value with the observed stadistic W2 and the p-value associated.

Author(s)

Veronica Paton Romero, Universidad Rey Juan Carlos, Spain <v.paton@alumnos.urjc.es>

References

Cabras and Castellanos (2008) Default Bayesian goodness-of-fit tests for the skew-normal model.

See Also

`W2.sn`

Examples

```
data(ais, package="sn")
attach(ais)
data=bmi
data(prior.lambda)
data(W2.quantiles)
W2.sn(data)
```

Index