Package 'MixedIndTests'

May 5, 2022

Type Package

Title Tests of Randomness and Tests of Independence

Version 0.9.1

Description Functions for testing randomness for a univariate time series with arbitrary distribution (discrete, continuous, mixture of both types) and for testing independence between random variables with arbitrary distributions. The test statistics are based on the multilinear empirical copula and multipliers are used to compute P-values. The test of independence between random variables appeared in Genest, Nešlehová, Rémillard & Murphy (2019).

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.5.0), doParallel, parallel, foreach, stats, copula

Imports ggplot2,survey

RoxygenNote 7.1.2

NeedsCompilation yes

Author Bouchra R. Nasri [aut, cre, cph],
Bruno N Remillard [aut],
Johanna G Neslehova [aut],
Christian Genest [aut]

Maintainer Bouchra R. Nasri <bouchra.nasri@umontreal.ca>

Repository CRAN

Date/Publication 2022-05-05 06:40:05 UTC

R topics documented:

`AutoDep` *Dependogram for Kendall's tau and Spearman's rho*

Description

This function, used in EstDepSerial, draws the P-values of Kendall's tau and Spearman's rho for a given number of lags.

Usage

```
AutoDep(out)
```

Arguments

`out`	List of the output of EstDepSerial (P-values, subsets)

References

B.R Nasri (2021). Tests of serial dependence for arbitrary distributions

Examples

```
out <-EstDepSerial(SimAR1Poisson(c(5,0.4),100),10)
AutoDep(out)
```

Dependogram *Dependogram for Cramer-von Mises statistics*

Description

This function, used in EstDep, TestIndCopula and TestIndSerCopula, draws the P-values of the Moebius Cramer-von Mises statistics from the multilinear copula and their combination for a tests of randomness for k consectives values X(1), ..., X(k) or for a test of independence between random variables.

Usage

```
Dependogram(out, stat = "CVM")
```

Arguments

out	List of the output from EstDep, TestIndCopula or TestIndSerCopula (P-values, subsets)
stat	Name of statistics to be used (default is "CVM")

References

Genest, Neslehova, Remillard & Murphy (2019). Testing for independence in arbitrary distributions

Examples

```
x <- matrix(rnorm(250),ncol=5)
out <-TestIndCopula(x)
Dependogram(out)
```

EstDep *Kendall's tau and Spearman's rho statistics for testing independence between random variables*

Description

This function computes the matrix of pairs of Kendall's tau and Spearman's rho statistics between random variables with arbitrary distributions.

Usage

```
EstDep(x, graph = FALSE)
```

Arguments

x	Data matrix
graph	Set to TRUE for a dependogram for all pairs of Kendall's taus and Spearman's rhos.

Value

stat	List of Kendall's tau and Spearman's rho statistics from multilinear copula, and test combinations LB
pvalue	P-values for the tests statistics

References

Genest, Neslehova, Remillard & Murphy (2018). Testing for independence in arbitrary distributions

Examples

```
x <- matrix(rnorm(500),ncol=10)
out <-EstDep(x)
```

EstDepSerial — *Kendall's tau and Spearman's rho statistics for testing randomness in a univariate time series*

Description

This function computes Kendall's tau and Spearman's rho statistics for tests of randomness in a time series with arbitrary distribution for pairs (X[i],X[i+k]), k=1:lags

Usage

```
EstDepSerial(x, lag, graph = FALSE)
```

Arguments

x	Time series
lag	Number of lags
graph	Set to TRUE for a dependogram for Kendall's tau and Spearman;s rho

Value

stat	List of Kendall's tau and Spearman's rho statistics from multilinear copula, and test combinations LB
pvalue	P-values for the tests statistics

References

B.R Nasri (2022). Tests of serial dependence for arbitrary distributions

Examples

```
out <-EstDepSerial(SimAR1Poisson(c(5,0.4),100),10)
```

`Finv` *Quantile function of margins*

Description

This function computes the quantile of seven cdf used in the simulatuons of Nasri (2022).

Usage

```
Finv(u, k)
```

Arguments

`u`	Vector of probabilities
`k`	Marginal distribution: [1] Bernoulli(0.8), [2] Poisson(6), [3] Negative binomial with r = 1.5, p = 0.2, [4] Zero-inflated Poisson (10) with w = 0.1 and P(6.67) otherwise, [5] Zero-inflated Gaussian, [6] Discretized Gaussian, [7] Discrete Pareto(1)

Value

`x`	Vector of quantiles

Author(s)

Bouchra R. Nasri January 2021

References

B.R Nasri (2022). Tests of serial dependence for arbitrary distributions

Examples

```
x = Finv(runif(100),2)
```

horseshoecrabs *Horseshoecrabs dataset*

Description

Horseshoe Crab Data from Table 3.2 of Agresti(2007). This data set consists of five variables, three of which are categorical, measured on 173 female crabs, each having a male attached in her nest.

Usage

```
data(horseshoecrabs)
```

Format

Data frame with 173 rows and 5 variables:

- X1: Color of the female (1: light medium, 2: medium, 3: dark medium, 4: dark)
- X2: Spine condition (1: both good. 2: one worn or broken, 3: both worn or broken)
- X3: Carapace width (cm)
- X4: Number of satellites, i.e., other males around the female
- X5: Weight (kg)

References

Agresti, A. (2007). An Introduction to Categorical data analysis, John Wiley & Sons, Wiley Series in Probability and Statistics, 2nd edition.

Examples

```
data(horseshoecrabs)
x =data.matrix(horseshoecrabs)
out = TestIndCopula(x,trunc.level=5,graph=TRUE)
```

lamb *Fetal lamb dataset*

Description

240 body movement measurements of a fetal lamb at consecutive 5 second intervals.

Usage

```
data(lamb)
```

Format

Count data.

References

Leroux B, Putterman M (1992). Maximum Penalized Likelihood estimation for independent and Markov-dependent Mixture models.Biometrics, 48, 545–558.

Examples

```
data(lamb)
plot(lamb)
```

---

select_p *Data-driven selection of p for the test of randomness*

Description

This function uses a AIC/BIC type criterion to select p based on the data.

Usage

```
select_p(X, p0 = 2, d = 5, q = 2.4, lambda = 0.25)
```

Arguments

X	Time series
p0	Minimum value of p (default is 2)
d	Maximum value of p (default is 5)
q	Constant for selecting between AIC and BIC type penalty (default is 2.4)
lambda	Penalty term (default is 0.25); small values lead to p=d, large value lead to p=p0

Value

p	Selected value of p

References

B.R Nasri (2021). Tests of serial dependence for arbitrary distributions

Examples

```
X <- SimAR1Poisson(c(5,0.2),100)
out <- select_p(X)
```

SimAR1Poisson *Simulation of a AR(1) Poisson process*

Description

Conditionally on the past, X[t] is Poisson with lambda[t] = a+bX[t-1]

Usage

```
SimAR1Poisson(param, n)
```

Arguments

param	Param[1] = a>0, param[2] = b, 0<=b <1 (for stationarity)
n	Length of the series.

Value

X	Simulated series

Examples

```
data <- SimAR1Poisson(c(5,0.4),500)
```

SimCopulaSeries *Simulation of a copula-based time series*

Description

This function simulates a Markovian time series (p-Markov for the Farlie-Gumbel-Morgenstern copula) with uniform margins using a copula family for the joint distribution of U[t], U[t-1].

Usage

```
SimCopulaSeries(family, n, tau = 0, param = NULL)
```

Arguments

family	"ind", "tent", "gaussian" , "t" , "clayton", "fgm", "frank", "gumbel", joe" , "plackett"
n	length of the time series
tau	Kendall's tau of the copula family
param	extra copula parameter: for "fgm", param in 2,3,... is the dimension of the copula; for "t", param = nu

Value

U	Simulated time series

Author(s)

Bouchra R. Nasri January 2021

References

B.R Nasri (2022). Tests of serial dependence for arbitrary distributions

Examples

```
U = SimCopulaSeries("fgm",100,0.2, 3) # for the FGM, |tau|<= 2/9
```

`TestIndCopula` *Statistics and P-values for a test of independence between random variables*

Description

This function computes Cramer-von Mises statistics and their combination for a tests of independence between random variables with arbitrary distributions. The P-values are computed using Gaussian multipliers.

Usage

```
TestIndCopula(
  x,
  trunc.level = 2,
  B = 1000,
  par = FALSE,
  ncores = 2,
  graph = FALSE
)
```

Arguments

x	Data matrix
trunc.level	Only subsets of cardinality <= trunc.level (default=2) are considered for the Moebius statistics.
B	Number of multipliers samples (default = 1000)
par	Set to TRUE if one prefers paraller computing (slower)
ncores	Number of cores for parallel computing (default is 2)
graph	Set to TRUE if one wants the dependogram of P-values for the Moebius statistics

Value

stat	List of Cramer-von Mises statistics cvm, Sn from the multilinear copula, and test combinations Tn and Tn2 (only pairs)
pvalue	Approximated P-values for the tests using Gaussian multipliers
card	Cardinaly of the subsets for the Moebius statistics
subsets	Subsets for the Moebius statistics

References

Genest, Neslehova, Remillard & Murphy (2019). Testing for independence in arbitrary distributions

Examples

```
x <- matrix(rnorm(250),ncol=5)
out <-TestIndCopula(x)
```

TestIndSerCopula *Statistics and P-values for a test of randomness for a univariate time series*

Description

This function computes Cramer-von Mises statistics from the multilinear copula and their combination for a tests of randomness for p consectives values X(1), ..., X(p). The p-values are computed using Gaussian multipliers.

Usage

```
TestIndSerCopula(
  x,
  p,
  trunc.level = 2,
  B = 1000,
  par = FALSE,
  ncores = 2,
  graph = FALSE
)
```

Arguments

x	Time series
p	Number of consecutive observations
trunc.level	Only subsets of cardinality <= trunc.level (default=2) are considered for the Moebius statistics.
B	Number of multipliers samples (default = 1000)

par	Set to TRUE if one prefers paraller computing (slower)
ncores	Number of cores for parallel computing (default = 2)
graph	Set to TRUE if one wants the dependogram of P-values for the Moebius statistics

Value

stat	List of Cramer-von Mises statistics cvm, Sn, and test combinations Tn and Tn2 (only pairs)
pvalue	Approximated P-values for the tests using Gaussian multipliers
card	Cardinaly of the subsets for the Moebius statistics
subsets	Subsets for the Moebius statistics

References

B.R Nasri (2022). Tests of serial dependence for arbitrary distributions

Examples

```
X <- SimAR1Poisson(c(5,0.2),100)
out <- TestIndSerCopula(X,5,3)
```

TestIndSerCopulaMulti *Statistics and P-values for a test of randomness for a multivariate time series*

Description

This function computes Cramer-von Mises statistics from the multilinear copula and their combination for a tests of randomness for p consecutives values of random vectors X(1), ..., X(p). The p-values are computed using Gaussian multipliers.

Usage

```
TestIndSerCopulaMulti(x, p, trunc.level = 2, B = 1000, graph = FALSE)
```

Arguments

x	Time series matrix
p	Number of consecutive vectors
trunc.level	Only subsets of cardinality <= trunc.level (default=2) are considered for the Moebius statistics.
B	Number of multipliers samples (default = 1000)
graph	Set to TRUE if one wants the dependogram of P-values for the Moebius statistics

Value

stat	List of Cramer-von Mises statistics cvm, tilde Sn, and test combinations tilde Tn and tilde Tn2 (only pairs), as defined in Nasri(2022).
pvalue	Approximated P-values for the tests using Gaussian multipliers

References

B.R Nasri (2022). Tests of serial dependence for arbitrary distributions

Examples

```
data(Y)
out <- TestIndSerCopulaMulti(Y,5,5)
```

X *AR(1) Poisson with parameters*

Description

Simulated AR(1) Poisson sequence of length n=100 with parameters c(5,0.4).

Usage

```
data(X)
```

Format

Count data.

Examples

```
data(X)
acf(X)
```

Xbin *Bernoulli sequence*

Description

Simulated Bernoulli sequence.

Usage

```
data(Xbin)
```

Format

Count data.

Examples

```
data(Xbin)
plot(Xbin)
```

Y *VAR(1) Poisson with parameters*

Description

Simulated VAR(1) Poisson sequence of length n=100.

Usage

```
data(Y)
```

Format

Count data.

Examples

```
data(Y)
acf(Y)
```

Index