Package: CopulaInference (via r-universe)

October 8, 2024

Type Package

Title Estimation and Goodness-of-Fit of Copula-Based Models with Arbitrary Distributions

Version 0.5.0

Description Estimation and goodness-of-fit functions for copula-based models of bivariate data with arbitrary distributions (discrete, continuous, mixture of both types). The copula families considered here are the Gaussian, Student, Clayton, Frank, Gumbel, Joe, Plackett, BB1, BB6, BB7,BB8, together with the following non-central squared copula families in Nasri (2020) <doi:10.1016/j.spl.2020.108704>: ncs-gaussian, ncs-clayton, ncs-gumbel, ncs-frank, ncs-joe, and ncs-plackett. For theoretical details, see, e.g., Nasri and Remillard (2023) <arXiv:2301.13408>.

License GPL-3

Encoding UTF-8

LazyData true

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

RoxygenNote 7.2.3

NeedsCompilation yes

Author Bouchra R. Nasri [aut, cre, cph], Bruno N Remillard [aut]

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

Date/Publication 2023-04-21 07:32:37 UTC

Repository https://bouchranasri.r-universe.dev

RemoteUrl https://github.com/cran/CopulaInference

RemoteRef HEAD

RemoteSha f107bafbcfc00c08180584b1245224f914be11a5

Contents

`AuxFun` *Auxiliary functions*

Description

This function computes the empirical margins, their left-limits, Kendall's tau and Spearman's rho for arbitrary data. Slower than AuxFunC based on C.

Usage

```
AuxFun(data)
```

Arguments

`data`	Matrix (x,y) of size n x 2

Value

`tau`	Kendall's tau
`rho`	Spearman's rho
`Fx`	Empirical cdf of x
`Fxm`	Left-limit of the empiricial cdf of x
`Fy`	Empirical cdf of y
`Fym`	Left-limit of the empiricial cdf of y

References

Nasri (2022). Test of serial dependence for arbitrary distributions. JMVA

Nasri & Remillard (2023). Tests of independence and randomness for arbitrary data using copula-based covariances, arXiv 2301.07267.

Examples

```
data(simgumbel)
out=AuxFun(simgumbel)
```

`AuxFunC` *Auxiliary functions using C*

Description

This function computes the empirical margins, their left-limits, Kendall's tau and Spearman's rho for arbitrary data

Usage

```
AuxFunC(data)
```

Arguments

`data`	Matrix (x,y) of size n x 2

Value

`tau`	Kendall's tau
`rho`	Spearman's rho
`Fx`	Empirical cdf of x
`Fxm`	Left-limit of the empirical cdf of x
`Fy`	Empirical cdf of y
`Fym`	Left-limit of the empirical cdf of y

References

Nasri (2022). Test of serial dependence for arbitrary distributions. JMVA

Nasri & Remillard (2023). Tests of independence and randomness for arbitrary data using copula-based covariances, arXiv 2301.07267.

Examples

```
data(simgumbel)
out=AuxFunC(simgumbel)
```

`BiEmpCdf` *Empirical bivariate cdf*

Description

This function computes the empirical joint cdf evaluated at all points (y1,y2)

Usage

```
BiEmpCdf(data, y1, y2)
```

Arguments

`data`	Matrix (x1,x2) of size n x 2
`y1`	Vector of size n1
`y2`	Vector of size n2

Value

`cdf`	Empirical cdf

Examples

```
data(simgumbel)
out=BiEmpCdf(simgumbel,c(0,1),c(-1,0,1))
```

CdfInv *Quantile function*

Description

This function computes the inverse of the cdf of a finite distribution for a vector of probabilities.

Usage

```
CdfInv(u, y, Fn)
```

Arguments

u	Vector of probabilities
y	Ordered values
Fn	Cdf

Value

x	Vector of quantiles

Examples

```
y=c(0,1,2)
Fn = c(0.5,0.85,1)
out=CdfInv(c(1:9)/10,y,Fn)
```

dncs *Density of non-central squared copula*

Description

This function computes the density of the non-central squared copula (ncs) associated with a one-parameter copula with parameter cpar, and parameters a1, a2 >0 .

Usage

```
dncs(data, family, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`family`	Copula family: "ncs-gaussian", "ncs-clayton", "ncs-frank", "ncs-gumbel", "ncs-joe", "ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	vector of copula parameter and non-centrality parameter a1,a2 >0

Value

`pdf`	Density

References

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
dncs(c(0.5,0.8),"ncs-clayton",par=c(2,1,2))
```

`dplac` *Density of Plackett copula*

Description

This function computes the density of the Plackett copula with parameter par>0.

Usage

```
dplac(data, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	Copula parameter >0

Value

`pdf`	Density

Examples

```
dplac(c(0.5,0.8),par=3,rotation=270)
```

EmpCdf *Empirical univariate cdf*

Description

This function computes the empirical cdf evaluated at all sample points

Usage

```
EmpCdf(x)
```

Arguments

`x`	Observations

Value

`Fx`	Empirical cdf
`Fxm`	Left limit of the empirical cdf
`Ix`	Indicator of atoms

data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAABIAAAASCAYAAABWzo5XAAAAWElEQVR42mNgGP

Examples

```
data(simgumbel)
out=EmpCdf(simgumbel[,1])
```

EstBiCop *Parameter estimation for bivariate copula-based models with arbitrary distributions*

Description

Computes the estimation of the parameters of a copula-based model with arbitrary distributions, i.e, possibly mixtures of discrete and continuous distributions. Parametric margins are allowed. The estimation is based on a pseudo-likelihood adapted to ties.

Usage

```
EstBiCop(
  data = NULL,
  family,
  rotation = 0,
  Fx = NULL,
  Fxm = NULL,
  Fy = NULL,
  Fym = NULL
)
```

Arguments

`data`	Matrix or data frame with 2 columns (X,Y). Can be pseudo-observations. If NULL, Fx and Fy must be provided.
`family`	Copula family: "gaussian", "t", "clayton", "frank", "gumbel", "joe", "plackett", "bb1", "bb6", "bb7","bb8","ncs-gaussian", "ncs-clayton", "ncs-gumbel", "ncs-frank", "ncs-joe","ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`Fx`	Marginal cdf function applied to X (default is NULL).
`Fxm`	Left-limit of marginal cdf function applied to X default is NULL).
`Fy`	Marginal cdf function applied to Y (default is NULL).
`Fym`	Left-limit of marginal cdf function applied to Y (default is NULL).

Value

`par`	Copula parameters
`family`	Copula family
`rotation`	Rotation value
`tauth`	Kendall's tau corresponding to the estimated parameter
`tauemp`	Empirical Kendall's tau (from the multilinear empirical copula)
`rhoSth`	Spearman's rho corresponding to the estimated parameter
`rhoSemp`	Empirical Spearman's tau (from the multilinear empirical copula)
`loglik`	Log-likelihood
`aic`	Aic value
`bic`	Bic value
`data`	Matrix of values (could be (Fx,Fy))
`F1`	Cdf of X (Fx if provided, empirical otherwise)
`F1m`	Left-limit of F1 (Fxm if provided, empirical otherwise)
`F2`	Cdf of Y (Fy if provided, empirical otherwise)
`F2m`	Left-limit of F2 (Fym if provided, empirical otherwise)
`ccdfx`	Conditional cdf of X given Y and it left limit

`ccdfxm`	Left-limit of ccdfx
`ccdfy`	Conditional cdf of Y given X and it left limit
`ccdfym`	Left-limit of ccdfy

References

Nasri & Remillard (2023). Identifiability and inference for copula-based semiparametric models for random vectors with arbitrary marginal distributions. arXiv 2301.13408.

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
set.seed(2)
data = matrix(rpois(20,1),ncol=2)
out0=EstBiCop(data,"gumbel")
```

`EstDep` *Kendall's tau and Spearman's rho*

Description

This function computes Kendall's tau and Spearman's rho for arbitrary data. These are invariant by increasing mappings.

Usage

```
EstDep(data)
```

Arguments

`data`	Matrix or data frame with 2 columns (X,Y). Can be pseudo-observations.

Value

`tau`	Kendall's tau
`rho`	Spearman's rho

References

Nasri (2022). Test of serial dependence for arbitrary distributions. JMVA

Nasri & Remillard (2023). Tests of independence and randomness for arbitrary data using copula-based covariances, arXiv 2301.07267.

Examples

```
data(simgumbel)
out=EstDep(simgumbel)
```

---

est_options — *Options for the estimation of the parameters of bivariate copula-based models*

Description

Sets starting values, upper and lower bounds for the parameters. The bounds are based on those in the rvinecopulib package.

Usage

```
est_options(family, tau = 0.5)
```

Arguments

`family`	Copula family: "gaussian", "t", "clayton", "frank", "gumbel", "joe", "plackett", "bb1", "bb6", "bb7","bb8","ncs-gaussian", "ncs-clayton", "ncs-gumbel", "ncs-frank", "ncs-joe","ncs-plackett".
`tau`	Estimated Kendall's tau to compute a starting point (default is 0.5)

Value

`LB`	Lower bound for the parameters
`UB`	Upper bound for the parameters
`start`	Starting point for the estimation

References

Nagler & Vatter (2002). rvinecopulib: High Performance Algorithms for Vine Copula Modeling. Version 0.6.2.1.3

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Nasri (2022). Test of serial dependence for arbitrary distributions. JMVA.

Nasri & Remillard (2023). Copula-based dependence measures for arbitrary data, arXiv 2301.07267.

Examples

```
out = est_options("bb8")
```

`Finv` *Quantile function of margins*

Description

This function computes the quantile of seven cdf used in 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(40),2)
```

`fnumber` *Family number corresponding to VineCopula package*

Description

Computes the number associated with a copula family (without rotation)

Usage

```
fnumber(family)
```

Arguments

`family`	Copula family: "gaussian", "t", "clayton", "frank", "gumbel", "joe", "plackett", "bb1", "bb6", "bb7","bb8".

Value

`fnumber`	Number

References

Nagler et al. (2023). VineCopula: Statistical Inference of Vine Copulas, version 2.4.5.

Examples

```
fnumber("bb1")
```

GofBiCop — *Goodness-of-fit for bivariate copula-based models with arbitrary distributions*

Description

Goodness-of-fit tests for copula-based models for data with arbitrary distributions. The tests statistics are the Cramer-von Mises statistic (Sn), the difference between the empirical Kendall's tau and the theoretical one, and the difference between the empirical Spearman's rho and the theoretical one.

Usage

```
GofBiCop(
  data = NULL,
  family,
  rotation = 0,
  Fx = NULL,
  Fxm = NULL,
  Fy = NULL,
  Fym = NULL,
  B = 100,
  n_cores = 1
)
```

Arguments

`data`	Matrix or data frame with 2 columns (X,Y). Can be pseudo-observations. If NULL, Fx and Fy must be provided.
`family`	Copula family: "gaussian", "t", "clayton", "frank", "gumbel", "joe", "plackett", "bb1", "bb6", "bb7","bb8","ncs-gaussian", "ncs-clayton", "ncs-gumbel", "ncs-frank", "ncs-joe","ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`Fx`	marginal cdf function applied to X (default is NULL).
`Fxm`	left limit of marginal cdf function applied to X default is NULL).
`Fy`	marginal cdf function applied to Y (default is NULL).
`Fym`	left limit of marginal cdf function applied to Y (default is NULL).
`B`	Number of bootstrap samples (default 100)
`n_cores`	Number of cores to be used for parallel computing (default is 1).

Value

`pvalueSn`	Pvalue of Sn in percent
`pvalueTn`	Pvalue of Tn in percent
`pvalueRn`	Pvalue of Rn in percent
`Sn`	Value of Cramer-von Mises statistic Sn
`Tn`	Value of Kendall's statistic Tn
`Rn`	Value of Spearman's statistic Rn
`cpar`	Copula parameters
`family`	Copula family
`rotation`	Rotation value
`tauth`	Kendall's tau (from the multilinear theoretical copula)
`tauemp`	Empirical Kendall's tau (from the multilinear empirical copula)
`rhoth`	Spearman's rho (from the multilinear theoretical copula)
`rhoemp`	Empirical Spearman's rho (from the multilinear empirical copula)
`parB`	Bootstrapped parameters
`loglik`	Log-likelihood
`aic`	AIC value
`bic`	BIC value

References

Nasri & Remillard (2023). Identifiability and inference for copula-based semiparametric models for random vectors with arbitrary marginal distributions. arXiv 2301.13408.

Nasri & Remillard (2023). Goodness-of-fit and bootstrapping for copula-based random vectors with arbitrary marginal distributions.

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
data = rvinecopulib::rbicop(10,"gumbel",rotation=0,2)
out=GofBiCop(data,family="gumbel",B=10)
```

`hncs` *Conditional distribution of non-central squared copula*

Description

This function computes the conditional distribution of the non-central squared copula (ncs) associated with a one-parameter copula with parameter cpar, and parameters a1, a2 >0 .

Usage

```
hncs(data, cond_var, family, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`cond_var`	Conditioning variable (1 or 2)
`family`	Copula family: "ncs-gaussian", "ncs-clayton", "ncs-frank", "ncs-gumbel", "ncs-joe", "ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	vector of copula parameter and non-centrality parameter a1,a2 >0

Value

`h`	Conditional cdf

References

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
hncs(c(0.5,0.8),1,"ncs-clayton",270,c(2,1,2))
```

hplac — *Conditional distribution of Plackett copula*

Description

This function computes the conditional distribution of the Plackett copula with parameter par>0.

Usage

```
hplac(data, cond_var, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`cond_var`	Conditioning variable (1 or 2)
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	Copula parameter >0

Value

`h`	Conditional cdf

Examples

```
hplac(c(0.5,0.8),1,270,3)
```

identifiability — *Identifiability of two-parameter copula families*

Description

Determines if a copula family is identifiable with respect to the empirical margins. One-parameter copula families ("gaussian","gumbel","clayton","frank","plackett","joe") are identifiable whatever the margins. The rank of the gradient of the copula on the range of the margins is evaluated at 10000 parameter points within the lower and upper bounds of the copula family.

Usage

```
identifiability(data = NULL, family, rotation = 0, Fx = NULL, Fy = NULL)
```

Arguments

`data`	Matrix or data frame with 2 columns (X,Y). Can be pseudo-observations. If NULL, Fx and Fy must be provided.
`family`	Copula family: "gaussian", "t", "clayton", "frank", "gumbel", "joe", "plackett”, "bb1", "bb6", "bb7","bb8","ncs-gaussian", "ncs-clayton", "ncs-gumbel", "ncs-frank", "ncs-joe","ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`Fx`	Marginal cdf function applied to X (default is NULL).
`Fy`	Marginal cdf function applied to Y (default is NULL).

Value

`out`	True or False

References

Nasri & Remillard (2023). Identifiability and inference for copula-based semiparametric models for random vectors with arbitrary marginal distributions. arXiv 2301.13408.

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
set.seed(1)
data = matrix(rpois(20,1),ncol=2)
out = identifiability(data,"gumbel")
```

`pncs` *Cdf for non-central squared copula*

Description

This function computes the distribution function of the non-central squared copula (ncs) associated a with one-parameter copula with parameter cpar, and parameters a1, a2 >0 .

Usage

```
pncs(data, family, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`family`	Copula family: "ncs-gaussian", "ncs-clayton", "ncs-frank", "ncs-gumbel", "ncs-joe", "ncs-plackett”.
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	vector of copula parameter and non-centrality parameter a1,a2 >0

Value

`cdf` Value of cdf

References

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
pncs(c(0.5,0.8),"ncs-clayton", par=c(2,1,2),rotation=270)
```

---

pplac *Cdf for Plackett copula*

Description

This function computes the distribution function of the Plackett copula with parameter par>0.

Usage

```
pplac(data, rotation = 0, par)
```

Arguments

`data`	Matrix (x,y) of size n x 2
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	Copula parameter >0

Value

`cdf` Value of cdf

Examples

```
pplac(c(0.5,0.8),270,3)
```

preparedata   *Computes unique values, cdf and pdf*

Description

This function computes the unique values, cdf and pdf for a series of data.

Usage

```
preparedata(x)
```

Arguments

`x`	Vector

Value

`values`	Unique (sorted) values
`m`	Number of unique values
`Fn`	Empirical cdf of the unique values
`fn`	Empirical pdf of the unique values

References

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

C. Genest, J.G. Neslehova, B.N. Remillard and O. Murphy (2019). Testing for independence in arbitrary distributions.

#' @examples x = c(0,0,0,2,3,1,3,1,2,0) out = prepare_data(x)

rhoplackett   *Spearman's rho for Plackett copula*

Description

Computes the theoretical Spearman's rho for Plackett copula

Usage

```
rhoplackett(cpar, rotation = 0)
```

Arguments

`cpar`	Copula parameter; can be a vector.
`rotation`	Rotation: 0 (default value), 90, 180, or 270.

Value

`rho`	Spearman's rho

References

Remillard (2013). Statistical Methods for Financial Engineering. CRC Press

Examples

```
rhoplackett(3,rotation=90)
```

---

rncs — *Simulation of non-central squared copula*

Description

This function computes generates a bivariate sample from a non-central squared copula (ncs) associated with a one-parameter copula with parameter cpar, and parameters a1, a2 >0 .

Usage

```
rncs(n, family, rotation = 0, par)
```

Arguments

`n`	Number of observations
`family`	Copula family: "ncs-gaussian", "ncs-clayton", "ncs-frank", "ncs-gumbel", "ncs-joe", "ncs-plackett".
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	vector of copula parameter and non-centrality parameter a1,a2 >0

Value

`U`	Observations

References

Nasri (2020). On non-central squared copulas. Statistics and Probability Letters.

Examples

```
rncs(100,"ncs-clayton",par=c(2,1,2))
```

`rplac` *Generates observations from the Plackett copula*

Description

This function generates observations from a Plackett copula with parameter par>0.

Usage

```
rplac(n, rotation = 0, par)
```

Arguments

`n`	Number of pairs to be generated
`rotation`	Rotation: 0 (default value), 90, 180, or 270.
`par`	Copula parameter >0

Value

`U`	Matrix of observations

Examples

```
rplac(10,rotation=90,par=2)
```

`simgumbel` *Simulated data*

Description

Simulated data from a Gumbel copula with parameter 2, Bernoulli margin for X1 and zero-inflated Gaussian margin for X2.

Usage

```
data(simgumbel)
```

Format

Data frame of numerical values

Examples

```
data(simgumbel)
plot(simgumbel,xlab="X1", ylab="X2")
```

`statcvm` *Goodness-of-fit statistics*

Description

Computation of goodness-of-fit statistics (Cramer-von Mises and the Kendall's tau)

Usage

```
statcvm(object)
```

Arguments

`object`	Object of class ‘EstBiCop‘.

Value

`Sn`	Cramer-von Mises statistic
`Tn`	Kendall's statistic
`Rn`	Spearman's statistic
`tauemp`	Empirical Kendall's tau
`tauth`	Kendall's tau of the multilineat theoretical copula
`rhoemp`	Empirical Spearman's rho
`rhoth`	Spearman's rho of the multilineat theoretical copula
`Y1`	Ordered observed values of X1
`F1`	Empirical cdf of Y1
`Y2`	Ordered observed values of X2
`F2`	Empirical cdf of Y2
`cpar`	Copula parameters
`family`	Copula family
`rotation`	Rotation value
`n`	Sample size

References

Nasri & Remillard (2023). Identifiability and inference for copula-based semiparametric models for random vectors with arbitrary marginal distributions. arXiv 2301.13408.

Examples

```
set.seed(2)
data = matrix(rpois(20,1),ncol=2)
out0 = EstBiCop(data,"gumbel")
out = statcvm(out0)
```

taucop — *Kendall's tau for a copula family*

Description

This function computes Kendall's tau for a copula family

Usage

```
taucop(family_number, cpar, rotation = 0)
```

Arguments

`family_number`	Integer from 1 to 10
`cpar`	Copula parameters
`rotation`	Rotation: 0 (default value), 90, 180, or 270.

Value

`tau`	Kendall's tau

Examples

```
taucop(4,2,270) # Gumbel copula
```

tauplackett — *Kendall's tau for Plackettfamily*

Description

This function computes Kendall's tau for Plackett family using numerical integration

Usage

```
tauplackett(cpar, rotation = 0)
```

Arguments

`cpar`	Copula parameter >0
`rotation`	Rotation: 0 (default value), 90, 180, or 270.

Value

`tau`	Kendall's tau

Examples

```
tauplackett(2,270)
```

Index