Package: satdad (via r-universe)

Type Package

Title Sensitivity Analysis Tools for Dependence and Asymptotic Dependence

Version 1.1

Date 2023-03-29

Maintainer Cécile Mercadier <mercadier@math.univ-lyon1.fr>

Description Tools for analyzing tail dependence in any sample or in particular theoretical models. The package uses only theoretical and non parametric methods, without inference. The primary goals of the package are to provide: (a)symmetric multivariate extreme value models in any dimension; theoretical and empirical indices to order tail dependence; theoretical and empirical graphical methods to visualize tail dependence.

License GPL (>= 3)

Imports Rcpp (>= 0.12.10)

LinkingTo Rcpp, RcppArmadillo, BH

RoxygenNote 7.2.3

Encoding UTF-8

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation yes

Depends igraph, maps, partitions, graphicalExtremes, R (>= 3.5.0)

LazyData true

Author Cécile Mercadier [aut, cre]

Date/Publication 2023-03-29 15:00:02 UTC

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

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

RemoteRef HEAD

RemoteSha fd80dca17bb79073920ec18721270259c7cdec10

Contents

---

copArchimaxMevlog	*cop-ell-psi-psiinv- functions for Archimax Mevlog models.*

Description

Copula function, stable tail dependence function, psi function, psi inverse function for Archimax Mevlog models.

Usage

```
copArchimaxMevlog(x, ds,  dist = "exp", dist.param = 1)
ellArchimaxMevlog(x, ds)
psiArchimaxMevlog(t, dist = "exp", dist.param = 1)
psiinvArchimaxMevlog(t, dist = "exp", dist.param = 1)
```

Arguments

x	A vector of size d or a (N.x times d) matrix.
ds	An object of class ds.
dist	The underlying distribution. A character string among "exp" (the default value), "gamma" and "ext".
dist.param	The parameter associated with the choice dist. If dist is "exp", then dist.param is a postive real, the parameter of an exponential distribution. The default value is 1. If dist is "gamma", then dist.param is a vector that concatenates the shape and scale parameters (in this order) of a gamma distribution.
t	A non negative scalar or vector.

Details

The tail dependence structure is set by a ds object. See Section **Value** in `gen.ds`.

Turning to Archimax structures, we follow Charpentier et al. (2014). Their algorithm (4.1 of p. 124) has been applied in `rArchimaxMevlog` to generate observations sampled from the copula

$$C(x_1, ..., x_d) = \psi(\ell(\psi^{-1}(x_1), ..., \psi^{-1}(x_d)))$$

when ℓ is here the stable tail dependence function of a Mevlog model. In this package, the stdf function ℓ is completely characterized by the ds object. See `ellMevlog`.

Value

When the underlying distribution `dist` is

- "exp" ; For a positive λ given by `dist.param`, $\psi(t) = \frac{\lambda}{t+\lambda}$ and $\psi^{-1}(t) = \lambda \frac{1-t}{t}$.
- "gamma" ; For positive scale σ and shape a given by `dist.param`, $\psi(t) = \frac{1}{(t+\sigma)^a}$ and $\psi^{-1}(t) = \frac{t^{-1/a}-1}{\sigma}$.
- "ext" ; $\psi(t) = \exp(-t)$ and $\psi^{-1}(t) = -\ln(t)$.

`copArchimaxMevlog` returns the copula function $C(x_1, ..., x_d) = \psi(\ell(\psi^{-1}(x_1), ..., \psi^{-1}(x_d)))$.

`ellArchimaxMevlog` returns the stable tail dependence function $\ell(x_1, ..., x_d)$.

`psiArchimaxMevlog` returns the psi function $\psi(t)$.

`psiinvArchimaxMevlog` returns the psi inverse function $\psi^{-1}(t)$.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Charpentier, A., Fougères, A.-L., Genest, C. and Nešlehová, J.G. (2014) Multivariate Archimax copulas. *Journal of Multivariate Analysis*, **126**, 118–136.

See Also

`rArchimaxMevlog`, `gen.ds`, `ellMevlog`

Examples

```
## Fix a 7-dimensional tail dependence structure
ds7 <- gen.ds(d = 7)

## Fix the parameters for the underlying distribution
(lambda <- runif(1, 0.01, 5))
(shape <- runif(1, 0.01, 5))
(scale <- runif(1, 0.01, 5))

## Fix x and t
x <- c(0.8, 0.9, 0.5, 0.8, 0.4, 0.9, 0.9)
t <- 2
```

```
## Evaluate the functions under the underlying exponential construction
copArchimaxMevlog(x = x, ds = ds7, dist = "exp", dist.param = lambda)
ellArchimaxMevlog(x = x, ds = ds7)
psiArchimaxMevlog(t = t, dist = "exp", dist.param = lambda)
psiinvArchimaxMevlog(t = t, dist = "exp", dist.param = lambda)

## Evaluate the functions under the underlying gamma construction
copArchimaxMevlog(x = x, ds = ds7, dist = "gamma", dist.param = c(shape, scale))
ellArchimaxMevlog(x = x, ds = ds7)
psiArchimaxMevlog(t = t, dist = "gamma", dist.param = c(shape, scale))
psiinvArchimaxMevlog(t = t, dist = "gamma", dist.param = c(shape, scale))
```

---

ec   *Extremal coefficients for Mevlog models.*

Description

Theoretical extremal coefficients for Mevlog models. A `Mevlog` model is a multivariate extreme value (symmetric or asymmetric) logistic model.

Usage

```
ec(ds, ind = 2, norm = FALSE)
```

Arguments

ds	An object of class `ds`.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is `ind = 2`, all pairwise coefficients are computed.
norm	A boolean. 'FALSE' (the default): ec is computed. 'TRUE': inverse normalized ec is computed.

Details

The tail dependence structure is set by a ds object. It thus corresponds to the stable tail dependence function ℓ. The way to deduce the stable tail dependence function ℓ from ds is explained in the Details section of gen.ds.

Value

The function returns a list of two elements:

- subsets A list of subsets from $\{1, ..., d\}$.

  When `ind` is given as an integer, `subsets` is the list of subsets from $\{1, ..., d\}$ with cardinality `ind`. When `ind` is the list, it corresponds to `subsets`.

When ind = "with.singletons" subsets is the list of all non empty subsets in $\{1, ..., d\}$.

When ind = "all" subsets is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- ec A vector of theoretical extremal coefficients associated with the list subsets.

  An extremal coefficient associated with the subset I is $\ell(1_I, 0_{I^c})$. Its value lies in $(1, |I|)$.

  When norm = TRUE, then inverse normalized ec are computed by $\dfrac{|I| - ec}{|I| - 1}$.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

Tiago de Oliveira, J. (1962/63) Structure theory of bivariate extremes, extensions. *Estudos de Matematica, Estatistica, e Economicos*, 7:165–195.

Smith, R. L. (1990) Max-stable processes and spatial extremes. *Dept. of Math., Univ. of Surrey*, Guildford GU2 5XH, England.

See Also

[ellMevlog](#), [gen.ds](#), [graphs](#), [tsic](#)

Examples

```
## Fix a 4-dimensional asymmetric tail dependence structure
ds4 <-  gen.ds(d = 4)
## Compute all theoretical extremal coefficients
ec(ds = ds4, ind = "with.singletons")
## Compute theoretical extremal coefficients associated with the support of ds4
ec(ds = ds4, ind = ds4$sub)

## Fix a 6-dimensional asymmetric tail dependence structure
ds6 <- gen.ds(d = 6, sub = list(1:2,2:5,5:6))
## Compute all theoretical extremal coefficients on subsets with cardinality 5
ec(ds = ds6, ind = 5)
## Compute inverse renormalized ec
ec(ds = ds6, ind = list(1:2,1:4,1:6), norm = TRUE)
```

---

ecEmp   *Empirical Extremal coefficients.*

Description

Computes on a sample the extremal coefficients associated with threshold k.

Usage

```
ecEmp(sample, ind = 2, k, norm = TRUE)
```

Arguments

sample	A (n times d) matrix.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = 2, all pairwise coefficients are computed.
k	An integer smaller or equal to n.
norm	A boolean. 'FALSE' (the default): empirical ec is computed. 'TRUE': inverse normalized empirical ec is computed.

Value

The function returns a list of two elements:

- subsets A list of subsets from $\{1, ..., d\}$.

  When ind is given as an integer, subsets is the list of subsets from $\{1, ..., d\}$ with cardinality ind. When ind is the list, it corresponds to subsets.

  When ind = $"$with.singletons$"$ subsets is the list of all non empty subsets in $\{1, ..., d\}$.

  When ind = $"$all$"$ subsets is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- ec A vector of empirical extremal coefficients.

  An empirical extremal coefficient associated with the subset I is $\hat{\ell}_{k,n}(1_I, 0_{I^c})$. Its value lies in $(1, |I|)$.

  When norm = TRUE, then inverse normalized empirical ec are computed by $1 - \dfrac{\hat{\ell}_{k,n}(1_I, 0_{I^c})}{|I|}$.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

See Also

ec, ellEmp, graphsEmp

Examples

```
## We produce below a figure on the dataset used in Mercadier  and Roustant (2019).

data(France)

ec_ymt <- ecEmp(sample = France$ymt, ind = 2, k = 25)

## The 9 largest inverse empirical pairwise extremal coefficients.
 graphsMapEmp(France$ymt, region='france', coord=France$coord, k=25, which="iecgraph", select=9)
```

```
## The 30 largest inverse empirical pairwise extremal coefficients.
graphsMapEmp(France$ymt, region='france', coord=France$coord, k=25, which="iecgraph", select=30)

## All the inverse empirical pairwise extremal coefficients.
graphsMapEmp(France$ymt, region='france', coord=France$coord, k=25, which="iecgraph")
```

---

ellEmp   *Empirical stable tail dependence function.*

Description

The stable tail dependence function of `sample` is estimated at each row of `x` and for all values of the threshold parameter `k`.

Usage

```
ellEmp(sample, x, k)
```

Arguments

sample	A (n times d) matrix.
x	A (N.x times d) matrix.
k	A vector of N.k integers smaller or equal to n.

Value

A (N.k times N.x) matrix is returned.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Huang, X. (1992). Statistics of bivariate extremes. PhD Thesis, Erasmus University Rotterdam, Tinbergen Institute Research series No. 22.

de Haan, L. and Resnick, S. I. (1993). Estimating the limit distribution of multivariate extremes. Communications in Statistics. Stochastic Models 9, 275–309.

Fougeres, A.-L., de Haan, L. and Mercadier, C. (2015). Bias correction in multivariate extremes. Annals of Statistics 43 (2), 903–934.

See Also

[ellMevlog](), [gen.ds]()

Examples

```
## Fix a 5-dimensional asymmetric tail dependence structure
ds5 <- gen.ds(d = 5)

## Construct a 1000-sample of Mevlog random vector associated with ds5
sample5 <- rMevlog(n = 1000, ds = ds5)

## Select 3 vectors in R^5
x5 <- matrix(runif(5*3), ncol = 5)

## Select 4 values for the threshold parameter
k5 <- (2:5)*10

## Estimation of the stable tail dependence function
# We thus get a 4 x 3 matrix
ellEmp(sample = sample5, x = x5, k = k5)

## Theoretical values of the stable tail dependence function inherited from ds5
ellMevlog(x = x5, ds = ds5)
```

France	*Dataset. Yearly Maxima of Temperature and coordinates of 21 French cities 1946-2000.*

Description

The France dataset is a list of two elements

- $ymt a data frame of 55 rows and 21 columns, constructed after extraction from www.ecad.eu. The value at row i and colum j is the yearly maximum of temperature for the year $1946 + i - 1$ in the $j - th$ French city.

- $coord a list of two elements: Latitude $lat and Longitude $lon of 21 French cities.

The name of the rows of $ymt are the year of the study: 1946–2000. The column names of $ymt are those of 21 French cities listed below.

[1] "MARSEILLE OBS. PALAIS-LONCHAMP" [2] "BOURGES AERODROME" [3] "BLAGNAC AEROP. TOULOUSE-BLAGNAC" [4] "MERIGNAC AEROPORT DE BORDEAUX" [5] "DE-OLS CHATEAUROUX AERODROME DE DEOLS" [6] "PERPIGNAN" [7] "BRON LYON AERO-PORT" [8] "PARIS-14E PARC MONTSOURIS" [9] "RENNES" [10] "STRASBOURG-ENTZHEIM" [11] "NANCY" [12] "ORLEANS" [13] "BESANCON" [14] "LA-ROCHELLE" [15] "BEAUVAIS-TILLE" [16] "LE MANS" [17] "METZ-FRESCATY" [18] "MONTELIMAR" [19] "NIMES" [20] "VICHY-CHARMEIL" [21] "COGNAC"

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Klein Tank, A.M.G. and Coauthors, (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. of Climatol.*, **22**, 1441–1453. Data and metadata available at www.ecad.eu

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

tsicEmp, ecEmp, graphsMapEmp

Examples

```
data(France)
maps::map('france',col='gray')
points(France$coord$lon,France$coord$lat, pch = 20, col = 1)
text(France$coord$lon,France$coord$lat+0.3,labels=1:21,cex=.8)
```

---

gen.ds	*Generate a Mevlog tail dependence structure.*

Description

The function gen.ds creates (possibly randomly) a tail dependence structure for a multivariate extreme value logistic (Mevlog) model.

Usage

```
gen.ds(d, type = "alog", sub = NULL, dep = NULL, asy = NULL, mnns = d)
```

Arguments

d	The dimension.
type	The type of the model; represented by a character string. This is similar to the option model of rmvevd. It must be either "log" or "alog" (the default), for the symmetric logistic and the asymmetric logistic model respectively.
sub	An optional list of subsets of $\{1, ..., d\}$ involved in the tail dependence structure. If type = "log", then sub should be given by $(1, \ldots, d)$, which is the way the code NULL will be interpreted. If type = "alog" and sub = NULL then a random list of vectors, subsets of $\{1, ..., d\}$, is created. The cardinality of non singleton subsets in sub is given by mnns. If the user provides sub, it has to be a list of vectors, subsets of $\{1, ..., d\}$, where each component from $\{1, ..., d\}$ appears at least once; Otherwise, one should add the missing singleton(s).
dep	An optional vector of dependence parameter(s). If type = "log", dep should be a single value. Otherwise, if type = "alog" and if the list sub is provided, then the length of the vector dep should be equal to that of the list sub (or a single value that will be replicated the length of sub times). Among these

values, the dependence parameters associated singletons have to be equal to one. Otherwise, the values of dep associated to singleton will be ignored (and set to one). When dep = NULL its values are randomly generated.

asy   An optional list of asymmetric weights. If type = "log", then asy should be the vector $(1, \ldots, 1)$, which is the way the code NULL will be interpreted. If type = "alog" and if sub is provided, the length of the list asy should be in accordance with the length of sub. If asy = NULL then the values are randomly generated. Note that asy satisfies the sum-to-one constraints.

mnns   The default value is arbitrarily equal to d. When sub = NULL, the list sub is randomly generated, and its size is closely related to mnns. The latter represents the number of non singletons subsets included in sub.

Details

A multivariate extreme value logistic (Mevlog) model is symmetric or asymmetric.

- type = "log". It generates a multivariate symmetric logistic model. Such model is a well-known generalization of the bivariate extreme value logistic model introduced by Gumbel (1960). The parameter 'dep' (with $0 < \text{'}dep\text{'} \leq 1$) is the only parameter needed to write the following equation

$$\ell(u) = (\sum_{i=1}^{d} u_i^{1/\text{dep}})^{\text{dep}}.$$

  If the parameter dep is missing, the function gen.ds will randomly generate its value from a standard uniform distribution. The list asy is reduced to a vector of ones whereas the list sub only contains the maximal vector $(1, \ldots, d)$.

  This is a special case of the multivariate asymmetric logistic model (alog case).

- type = "alog". It generates a multivariate asymmetric logistic model, which has been first introduced by Tawn (1990). We have

$$\ell(u) = \sum_{b \in B} (\sum_{i \in b} (\beta_{i,b} u_i)^{1/\alpha_b})^{\alpha_b}$$

  where B is the power set of $\{1, ..., d\}$ (or a strict subset of the power set), the dependence parameters α_b lie in $(0, 1]$ and the collection of asymmetric weights $\beta_{i,b}$ are coefficients from [0,1] satisfying $\forall i \in \{1, \ldots, d\}, \sum_{b \in B: i \in b} \beta_{i,b} = 1$. Missing asymmetric weights $\beta_{i,b}$ are assumed to be zero.

The function gen.ds generates here an object of class ds which corresponds in this package to the stable tail dependence function ℓ. The class ds consists of:

- the dimension d.

- the type "log" or alog.

- the list sub that corresponds to B. When sub is provided, the same list of subsets is returned, eventually sorted. When sub = NULL then sub is a list of subsets of the power set of $\{1, ..., d\}$. When the option mnns is used, the latter integer indicates the cardinality of non singleton subsets in B. the dependence parameter dep or the vector of dependence parameters dep. When missing, these coefficients are obtained from independent standard uniform sampling.

the list `asy` of asymmetric weights $\beta_{i,b}$ for $b \in B$ and $i \in b$. When missing, these coefficients are obtained from independent standard uniform sampling followed by renormalization in order to satisfy the sum-to-one constraints.

Value

`gen.ds` returns an object representing a tail dependence structure for Mevlog models. Such object is a list containing the following components:

- `d` The dimension.
- `type` The type of the model either `"log"` or `"alog"`.
- `sub` The list of subsets of $\{1, ..., d\}$ involved in the tail dependence support.
- `dep` The vector of dependence parameter(s).
- `asy` The list of asymmetric weights.

Note

The first interest of the `gen.ds` function is to generate randomly a tail dependence structure. Since `sub` and `asy` become quickly very large lists as d increases, it is very convenient to obtain automatically well-defined tail dependence structures for multivariate extreme value logistic models.

The second interest of the `gen.ds` function is to produce partial models where all subsets do not necessarily contribute to the tail dependence support.

The function `gen.ds` does not manage margins characteristics which will be handle by the option `mar` in the `r-d-p-Mevlog` functions.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Gumbel, E. J. (1960) Distributions des valeurs extremes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris*, **9**, 171–173.

Stephenson, A. (2002) evd: Extreme Value Distributions. *R News*, 2(2):31–32.

Tawn, J. A. (1990) Modelling multivariate extreme value distributions. *Biometrika*, **77**, 245–253.

See Also

`ellMevlog`, `graphs`

Examples

```
## Fix a 5-dimensional symmetric tail dependence structure
## The dependence paramater is fixed to .7
(ds5 <- gen.ds(d = 5, dep = .7, type = "log"))

## Fix a 3-dimensional asymmetric tail dependence structure
## The list sub and asy are provided ; The vector dep is randomly generated
```

```
(ds3 <- gen.ds(d = 3, sub = list(c(1,2), c(1,2,3)), asy = list(c(0.4,0.6), c(0.6,0.4,1))))
graphs(ds = ds3)

## Fix a 8-dimensional asymmetric tail dependence structure
## The lists sub and asy, as the vector dep, are randomly generated
(ds8 <- gen.ds(d = 8))
graphs(ds = ds8)
```

---

graphs   *Graphs of the tail dependence structure for Mevlog models.*

Description

Tail dependograph and Inverse extremal coefficients graph for Mevlog models. A `Mevlog` model is
a multivariate extreme value (symmetric or asymmetric) logistic model.

Usage

```
graphs(
  ds,
  names = NULL,
  n.MC = 1000,
  which = "taildependograph",
  random = FALSE,
  thick.td = 5,
  thick.ec = 5
)
```

Arguments

ds	An object of class ds.
names	A character vector of length d which replaces as.character(1:d) (the default ones).
n.MC	Monte Carlo sample size. Default value is 1000. See details in `tsic`.
which	A character string: taildependograph (the default), iecgraph, or both,
random	A boolean. 'FALSE' (the default): the vertex positions are fixed along a circle. 'TRUE': some randomness is applied for positioning the vertices.
thick.td	A numeric value for the maximal thickness of edges in taildependograph. Default value is 5.
thick.ec	A numeric value for the maximal thickness of edges in iecgraph. Default value is 5.

Details

The tail dependence structure is set by a ds object. It thus corresponds to the stable tail dependence
function ℓ. The way to deduce the stable tail dependence function ℓ from ds is explained in the
Details section of `gen.ds`.

Value

The function returns either the tail dependograph or the inverse extremal coefficients graph, or both, for the tail dependence structure 'ds'.

The tail dependograph displays pairwise tail superset importance coefficients, which measure the extent to which pairs of components (and their supersets) contribute to the overall variance of the stable tail dependence function. We refer to Mercadier, C. and Roustant, O. (2019) for more details. These coefficients are computed using the 'tsic' function with the '"ind = 2"' option.

The inverse extremal coefficients graph shows the inverse renormalized pairwise coefficients computed as $\theta_{ij} = 1 - \ell(1_i, 1_j, \mathbf{0})/2$.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

Tiago de Oliveira, J. (1962/63) Structure theory of bivariate extremes, extensions. *Estudos de Matematica, Estatistica, e Economicos*, 7:165–195.

Smith, R. L. (1990) Max-stable processes and spatial extremes. *Dept. of Math., Univ. of Surrey*, Guildford GU2 5XH, England.

See Also

tsic, ec, ellMevlog

Examples

```
## Fix a 8-dimensional asymmetric tail dependence structure
ds8 <- gen.ds(d = 8)

## Plot the graphs that illustrate  characteristics of the tail dependence structure
graphs(ds = ds8, which = "both")
```

---

graphsEmp   *Empirical graphs of the tail dependence structure.*

Description

Empirical tail dependograph and empirical inverse extremal coefficients graph of the tail dependence structure on a `sample` associated with threshold k.

Usage

```
graphsEmp(
  sample,
  layout = NULL,
  names = NULL,
  k,
  which = "taildependograph",
  select = NULL,
  simplify = FALSE,
  random = FALSE,
  thick.td = 5,
  thick.ec = 5
)
```

Arguments

sample	A (n times d) matrix.
layout	The vertex coordinates as a (d times 2) matrix. The default is NULL. See also the parameter random.
names	A character vector of length d which replaces as.character(1:d) (the default ones).
k	An integer smaller or equal to n.
which	A character string: taildependograph (the default), iecgraph or both.
select	If select = NULL (default) all edges are plotted. If select is an integer between 1 and the number of possible pairs of components of sample, then only the select largest edges are plotted.
simplify	If select is not NULL, and if a vertex is not associated with one of the selected edges, this vertex is not printed.
random	A boolean. 'FALSE' (the default): the vertex positions are fixed along a circle when layout is NULL. 'TRUE': some randomness is applied for positioning the vertices.
thick.td	A numeric value for the maximal thickness of edges in taildependograph. Default value is 5.
thick.ec	A numeric value for the maximal thickness of edges in iecgraph. Default value is 5.

Value

It returns both (or one among) the empirical tail dependograph and the empirical inverse extremal coefficients graph of the sample.

The empirical tail dependograph represents the pairwise empirical tail superset importance coefficients, see Mercadier, C. and Roustant, O. (2019). These indices are computed by the function tsicEmp. It measures how much a pair of components (included supersets of this pair of components) is involved in the asymptotic dependence of the sample.

The empirical Inverse extremal coefficients graph represents empirical pairwise coefficients that estimate $1 - \ell(1_i, 1_j, \mathbf{0})/2$.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

`ecEmp`, `tsicEmp`

Examples

```
## Fix a 8-dimensional asymmetric tail dependence structure
ds8 <- gen.ds(d = 8)

## Generate a 200-sample of Frechet margins Mevlog model associated with ds8
sample8 <- rMevlog(n = 200 , ds = ds8)

## Plot the tail dependograph of ds8
graphs(ds = ds8)

## Its empirical version for k = 20
graphsEmp(sample = sample8, k = 20)

## Its empirical version for k = 20 restricted to the 3 largest edges
graphsEmp(sample = sample8, k = 20, select = 3)

## Plot the Inverse extremal coefficients graph of ds8
graphs(ds = ds8, which = "iecgraph")

## Its empirical version for k = 20
graphsEmp(sample = sample8, k = 20, which = "iecgraph")

## Its empirical version for k = 20 restricted to the 3 largest edges
graphsEmp(sample = sample8, k = 20, which = "iecgraph", select = 3)

## Plot the empirical tail dependograph
## on river discharge data for tributaries
## of the Danube extracted from
##  Asadi P., Davison A.C., Engelke S. (2015).
## "Extremes on river networks."
## The Annals of Applied Statistics, 9(4), 2023 - 2050.
#NOT RUN dan <- graphicalExtremes::danube$data_clustered
#NOT loc <- as.matrix(graphicalExtremes::danube$info[,c('PlotCoordX', 'PlotCoordY')])
#NOT graphsEmp(dan,  k=50, layout = loc)
```

---

graphsMapEmp   *Empirical graphs drawn on geographical maps of the tail dependence*
                                structure.

Description

Empirical tail dependograph and Empirical inverse extremal coefficients graph drawn on geographical maps for the tail dependence structure of sample associated with threshold k.

Usage

```
graphsMapEmp(
  sample,
  k,
  which = "taildependograph",
  names = NULL,
  coord,
  region = NULL,
  select = NULL,
  thick.td = 5,
  thick.ec = 5,
  eps = 0.03
)
```

Arguments

sample	A (n times d) matrix.
k	An integer smaller or equal to n.
which	A character string: both, taildependograph (the default) or iecgraph.
names	A character vector for sample columns which replaces as.character(1:d) (the default ones).
coord	Latitudes and Longitudes associated with sample columns associated to region map when region is furnished.
region	A geographical region from maps package. The default value is NULL.
select	If select is NULL (the default) all edges are plotted. If select is an integer between 1 and the number of possible pairs of components of sample, then only the select largest edges are plotted.
thick.td	A numeric value for the maximal thickness of edges in taildependograph. Default value is 5.
thick.ec	A numeric value for the maximal thickness of edges in iecgraph. Default value is 5.
eps	A numerical graphical value fixing the distance between the plotted point and its names. The default value is 0.03.

Value

It returns both (or one among) the empirical tail dependograph and the empirical inverse extremal coefficients graph on a geographical map of the `sample`.

The empirical tail dependograph on a geographical map represents the pairwise empirical tail superset importance coefficients of the locations associated with `sample` columns, see Mercadier, C. and Roustant, O. (2019). These indices are computed by the function `tsicEmp`. It measures how much a pair of components (included supersets of this pair of components) is involved in the asymptotic dependence of the sample.

The empirical inverse extremal coefficients graph on a geographical map represents empirical pairwise coefficients of the locations associated with `sample` columns that estimate $1 - \ell(1_i, 1_j, \mathbf{0})/2$.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

Becker, R. A., Wilks, A. R. (Original S code), Brownrigg, R. (R version), Minka, T. P. and Deckmyn A. (Enhancements). (2022) maps : Draw Geographical Maps. *R package version 3.4.1.*

See Also

graphsEmp

Examples

```
data(France)

## Figure 9 (a) of Mercadier  and Roustant (2019).
graphsMapEmp(France$ymt, k = 55,
 coord = France$coord,  region = 'France', select = 9)

## Figure 9 (b) of Mercadier  and Roustant (2019).
graphsMapEmp(France$ymt, k = 55,
 coord = France$coord,  region = 'France', select = 30)

## Figure 9 (c) of Mercadier  and Roustant (2019).
graphsMapEmp(France$ymt, k = 55,
 coord = France$coord,  region = 'France')
```

---

plotClev   *Cleveland's Dot Plots of the tail dependence structure.*

Description

Global comparison of the theoretical tail superset importance coefficients (tsic) via a Cleveland's Dot Plot.

Usage

```
plotClev(ds, ind = "all", which = "tsic", labels = TRUE)
```

Arguments

ds	An object of class ds.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = "all".
which	A character string among "tsic" (normalized tsic plot), and "ec" (normalized ec plot).
labels	A boolean. 'TRUE' the default indicates that the names of the subsets are printed. 'FALSE' if only points are drawn.

Value

Draws a Cleveland dot plot of the normalized theoretical tsic when superset = TRUE, the default value.

Otherwise theoretical normalized ec are drawn.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

See Also

[plotClevEmp](#)

Examples

```
## Fix a 6-dimensional asymmetric tail dependence structure
## Two blocks of components are specified
ds6 <- gen.ds(d = 6, sub = list(1:4,5:6))

## Plot the associated Cleveland dot plot
plotClev(ds6)
```

---

plotClevEmp   *Empirical Cleveland's Dot Plots of the tail dependence structure.*

Description

Global comparison of the empirical tail superset importance coefficients (tsicEmp) via a Cleveland's Dot Plot.

Usage

```
plotClevEmp(sample, k, ind = "all", which = "tsic", labels = TRUE)
```

Arguments

sample	A (n times d) matrix.
k	An integer smaller or equal to n.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = "all".
which	A character string among "tsic" (empirical normalized tsic plot), and "ec" (empirical normalized ec plot).
labels	A boolean. 'TRUE' the default indicates that the names of the subsets are printed. 'FALSE' if only points are drawn.

Value

Draws a Cleveland dot plot of the normalized empirical tsic when superset = TRUE, the default value.

Otherwise empirical normalized ec are drawn.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

See Also

[plotClev](#)

Examples

```
## Fix a 5-dimensional asymmetric tail dependence structure
(ds5 <- gen.ds(d = 5))

## Generate a 1000-sample of Mevlog random vectors associated with ds5
sample5 <- rMevlog(n = 1000, ds = ds5)

## Plot the empirical Cleveland dot plot (restricted to pairs)
```

```
plotClevEmp(sample5,  k = 100,  ind = 2)
```

---

rArchimaxMevlog   *r function for Archimax Mevlog models.*

Description

Random vectors generation for some Archimax Mevlog models.

Usage

```
rArchimaxMevlog(n, ds, dist = "exp", dist.param = 1)
```

Arguments

n	The number of observations.
ds	An object of class ds.
dist	The underlying distribution. A character string among "exp" (the default value), "gamma" and "ext".
dist.param	The parameter associated with the choice dist. If dist is "exp", then dist.param is a postive real, the parameter of an exponential distribution. The default value is 1. If dist is "gamma", then dist.param is a vector that concatenates the shape and scale parameters (in this order) of a gamma distribution.

Details

We follow below Algorithm 4.1 of p. 124 in Charpentier et al. (2014). Let ψ defined by $\psi(x) = \int_0^\infty \exp(-xt)dF(t)$, the Laplace transform of a positive random variable with cumulative distribution function F.

Define the random vector $(U_1, ..., U_d)$ as $U_i = \psi(-\log(Y_i)/V)$ where

- Z has a multivariate extreme value distribution with stable tail dependence function ℓ ; here Z has standard Frechet margins,
- $(Y_1, ..., Y_d) = (\exp(-1/Z_1), ..., \exp(-1/Z_d))$ the margin transform of Z so that Y is sampled from the extreme value copula associated with ℓ,
- V has the distribution function F,
- Y and V are independent.

Then, U is sampled from the Archimax copula $C(u_1, ..., u_d) = \psi(\ell(\psi^{-1}(u_1), ..., \psi^{-1}(u_d)))$.

We restrict here the function ℓ to those associated with Mevlog models. See ellMevlog and gen.ds.

We restrict also the distribution of V to

- exponential ; For a positive λ, set $dF(t) = \lambda \exp(-\lambda t)1_{t>0}dt$, then $\psi(x) = \frac{\lambda}{x+\lambda}$ and $\psi^{-1}(x) = \lambda\frac{1-x}{x}$.
- gamma ; For positive scale σ and shape a, set $dF(t) = \frac{1}{\sigma^a\Gamma(a)}t^{a-1}\exp(-t/\sigma)1_{t>0}$, then $\psi(x) = \frac{1}{(x+\sigma)^a}$ and $\psi^{-1}(x) = \frac{x^{-1/a}-1}{\sigma}$.

Value

returns a `n x d` matrix containing `n` realizations of a d-variate Archimax Mevlog random vector.

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Charpentier, A., Fougères, A.-L., Genest, C. and Nešlehová, J.G. (2014) Multivariate Archimax copulas. *Journal of Multivariate Analysis*, **126**, 118–136.

See Also

`rMevlog`, `copArchimaxMevlog`, `psiArchimaxMevlog`, `psiinvArchimaxMevlog`, `gen.ds`

Examples

```
## Fix a  5-dimensional asymmetric tail dependence structure
(ds5 <- gen.ds(d = 5))

## Generate a 1000-sample of Archimax Mevlog random vectors
## associated with ds5 and underlying distribution gamma
(shape5 <- runif(1, 0.01, 5))
(scale5 <- runif(1, 0.01, 5))
sample5.gamma <- rArchimaxMevlog(n = 1000, ds = ds5, dist = "gamma", dist.param = c(shape5, scale5))

## Compare theoretical (left) and empirical (right) tail dependographs
oldpar <- par(mfrow = c(1,2))
graphs(ds = ds5)
graphsEmp(sample = sample5.gamma, k = 100)
par(oldpar)

## Generate a 1000-sample of Archimax Mevlog random vectors
## associated with ds5 and underlying distribution exp
(lambda <- runif(1, 0.01, 5))
sample5.exp <- rArchimaxMevlog(n = 1000, ds = ds5, dist = "exp", dist.param = lambda)
## Compare theoretical (left) and empirical (right) tail dependographs
graphs(ds = ds5)
graphsEmp(sample = sample5.exp, k = 100)
```

---

rMevlog   *r-p-d-ell- functions for Mevlog models.*

Description

Random vectors generation (`rMevlog`), cumulative distribution function (`pMevlog`), probability density function (`dMevlog`), stable tail dependence function (`ellMevlog`) for Mevlog models. A `Mevlog` model is a multivariate extreme value (symmetric or asymmetric) logistic model.

Usage

```
rMevlog(n, ds, mar = c(1,1,1))
pMevlog(x, ds, mar = c(1,1,1))
dMevlog(x, ds, mar = c(1,1,1))
ellMevlog(x, ds)
```

Arguments

n	The number of observations.
ds	An object of class ds.
mar	A vector of length 3 or a (d times 3) matrix. See details.
x	A vector of size d or a matrix with d columns.

Details

The tail dependence structure is set by a ds object. See Section **Value** in `gen.ds`.

The marginal information mar is given by a 3-dimensional vector (the order should be location, scale and shape) or a matrix with 3 columns depending on whether the components share the same characteristics or not. When the marginal parameters differ, mar is a matrix containing d locations in the first column, d scales in the second column and d shapes in the third column.

The (a)symmetric logistic models respectively are simulated in 'rMevlog' using Algorithms 2.1 and 2.2 in Stephenson(2003).

Value

rMevlog returns a (n times d) matrix containing n realizations of a d-variate Mevlog random vector with margins mar and tail dependence structure ds.

pMevlog returns a scalar (when x is a numeric vector) or a vector (when x is a numeric matrix, in which case the evaluation is done across the rows). The margins are provided by mar and the tail dependence structure through a ds object.

dMevlog returns a scalar (when x is a numeric vector) or a vector (when x is a numeric matrix, in which case the evaluation is done across the rows). The margins are provided by mar and the tail dependence structure through a ds object.

ellMevlog returns a scalar (when x is a numeric vector) or a vector (when x is a numeric matrix, in which case the evaluation is done across the rows). The tail dependence structure is provided by a ds object.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Gumbel, E. J. (1960) Distributions des valeurs extremes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris*, **9**, 171–173.

Stephenson, A. (2002) evd: Extreme Value Distributions. *R News*, 2(2):31–32.

Stephenson, A. (2003) Simulating Multivariate Extreme Value Distributions of Logistic Type. *Extremes*, **6**, 49–59.

Tawn, J. A. (1990) Modelling multivariate extreme value distributions. *Biometrika*, **77**, 245–253.

See Also

[gen.ds](), [tsic](), [ec](), [graphs]()

Examples

```
## Fix a 3-dimensional symmetric tail dependence structure
ds3 <- gen.ds(d = 3, type = "log")

## The dependence parameter is given by
ds3$dep

## Generate a 1000-sample of Mevlog random vectors associated with ds3
## The margins are kept as standard  Frechet
sample3 <- rMevlog(n = 1000, ds = ds3)

## Fix a 10-dimensional asymmetric tail dependence structure
# The option \cdoe{mns = 4} produces a support involving subsets of cardinality 4 plus singletons.
ds10 <- gen.ds(d = 10,  mnns = 4)
## Margins differ from one to another
mar10 <- matrix(runif(10*3), ncol = 3)

## Generate a 50-sample of Mevlog random vectors associated with ds10 and mar10
sample10 <- rMevlog(n = 50, ds = ds10, mar = mar10)

## Continuing with ds3 ; we compute other attributes
## The cumulative distribution function
pMevlog(x = rep(1,3), ds = ds3)
# should be similar to :
# evd::pmvevd(q = rep(1,3), dep = ds3$dep, model = "log", d = 3, mar = c(1,1,1))
## The probability density function:
dMevlog(x = rep(1,3), ds = ds3, mar = c(1.2,1,0.5))
# should be similar to :
# evd::dmvevd(x = rep(1,3), dep = ds3$dep, model = "log", d = 3, mar = c(1.2,1,0.5))
## The stable tail dependence function:
ellMevlog(x = rep(1,3), ds = ds3)
```

Stock	Dataset. *Yearly maxima of Log Returns of ten stock indices 1990-2015.*

Description

This dataset consists of a matrix with years as rows and columns as stock indices. They appear in the following order: "SP500", "DJ", "NASD", "SMI", "EURS", "CAC", "DAX", "HSI", "SSEC", "NIKK". A cell gives the yearly maximum of Log Returns of the associated stock indices. The latter values have been extracted from the R package qrmdata of Hofert, M., Hornik, K. and McNeil, A.J. (2019).

Author(s)

Cécile Mercadier (`mercadier@math.univ-lyon1.fr`)

References

Hofert, M., Hornik, K. and McNeil, A.J. (2019). qrmdata: Data Sets for Quantitative Risk Management Practice. R package version 2019-12-03-1 URL https://CRAN.R-project.org/package=qrmdata.

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

graphsEmp

Examples

```
data(Stock)

## We reproduce below Figure 7(a) of Mercadier and Roustant (2019).

graphsEmp(Stock, k = 26, which = "taildependograph", names = colnames(Stock))

## We reproduce below Figure 8(a) of Mercadier and Roustant (2019).

graphsEmp(Stock, k = 26, which = "taildependograph", names = colnames(Stock), select = 9)

## We reproduce below Figure 8(b) of Mercadier and Roustant (2019).

graphsEmp(Stock, k = 26, which = "taildependograph", names = colnames(Stock), select = 20)
```

---

tic | *Tail importance coefficients for Mevlog models.*

Description

Computes the tail importance coefficients (tic) on a `Mevlog` model which is a multivariate extreme value (symmetric or asymmetric) logistic model, descibed here by its dependence structure.

Usage

```
tic(ds, ind = 2, n.MC = 1000, sobol = FALSE)
```

Arguments

ds	An object of class ds.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = 2, all pairwise coefficients are computed.
n.MC	Monte Carlo sample size. Default value is 1000. See details in tsic.
sobol	A boolean. 'FALSE' (the default). If 'TRUE': the index is normalized by the theoretical global variance.

Details

The tail dependence structure is specified using a ds object, which corresponds to the stable tail dependence function ℓ. The process for deducing the stable tail dependence function ℓ from ds is explained in the Details section of gen.ds.

The theoretical functional decomposition of the variance of the stdf ℓ consists in writing $D(\ell) = \sum_{I \subseteq \{1,...,d\}} D_I(\ell)$ where $D_I(\ell)$ measures the variance of $\ell_I(U_I)$ the term associated with subset I in the Hoeffding-Sobol decomposition of ℓ ; note that U_I represents a random vector with independent standard uniform entries. The theoretical tail importance coefficient (tic) is thus $D_I(\ell)$ and its sobol version is $S_I(\ell) = \dfrac{D_I(\ell)}{D(\ell)}$.

The function tic uses the Mobius inversion formula, see Formula (8) in Liu and Owen (2006), to derive the tic from the tsic. The latter are the tail superset importance coefficients obtained by the function tsic.

Value

The function returns a list of two elements:

- subsets A list of subsets from $\{1, ..., d\}$.

  When ind is given as an integer, subsets is the list of subsets from $\{1, ..., d\}$ with cardinality ind.

  When ind is a list, it corresponds to subsets.

  When ind = "with.singletons" subsets is the list of all non empty subsets in $\{1, ..., d\}$.

  When ind = "all" subsets is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- tic A vector of tail importance coefficients, or their Sobol versions when sobol = "TRUE".

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Liu, R. and Owen, A. B. (2006) Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc.*, **101(474)**:712–721.

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

tsic, ticEmp

ticEmp and tsic

Examples

```
## Fix a 4-dimensional asymmetric tail dependence structure
ds4 <- gen.ds(d = 4, sub = list(1:2,3:4,1:3))

## Compute all tic values
res4 <- tic(ds4, ind = "with.singletons", sobol = TRUE)

## Check the sum-to-one constraint of tail Sobol indices
sum(res4$tic)
```

---

ticEmp   *Empirical tail importance coefficients.*

Description

Computes on a sample the tail importance coefficients (tic) associated with threshold k. The value
may be renormalized by the empirical global variance (Sobol version).

Usage

```
ticEmp(sample, ind = 2, k, sobol = FALSE)
```

Arguments

sample	A (n times d) matrix.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = 2, all pairwise coefficients are computed.
k	An integer smaller or equal to n.
sobol	A boolean. 'FALSE' (the default). If 'TRUE': the index is normalized by the empirical global variance.

Details

The theoretical functional decomposition of the variance of the stdf ℓ consists in writing $D(\ell) = \sum_{I \subseteq \{1,...,d\}} D_I(\ell)$ where $D_I(\ell)$ measures the variance of $\ell_I(U_I)$ the term associated with subset I in the Hoeffding-Sobol decomposition of ℓ ; note that U_I represents a random vector with independent standard uniform entries. The theoretical tail variance contribution is thus $D_I(\ell)$ and the theoretical tail sobol index is $S_I(\ell) = \dfrac{D_I(\ell)}{D(\ell)}$.

Here, the function `ticEmp` evaluates $\hat{D}_{I,k,n}$ the empirical counterpart of $D_I(\ell)$ under the option `sobol = FALSE`, and $\hat{S}_{I,k,n}$ the empirical counterpart of $S_I(\ell)$ under the option `sobol = TRUE`.

Proposition 1 and Theorem 2 of Mercadier and Roustant (2019) furnish their rank-based expressions. For the subset of components I,

$$\hat{D}_{I,k,n} = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t \in I} (\min(\overline{R}_s^{(t)}, \overline{R}_{s'}^{(t)}) - \overline{R}_s^{(t)} \overline{R}_{s'}^{(t)}) \prod_{t \notin I} \overline{R}_s^{(t)} \overline{R}_{s'}^{(t)}$$

$$\hat{D}_{k,n} = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t \in I} \min(\overline{R}_s^{(t)}, \overline{R}_{s'}^{(t)}) - \prod_{t \in I} \overline{R}_s^{(t)} \overline{R}_{s'}^{(t)}$$

and $\hat{S}_{I,k,n} = \dfrac{\hat{D}_{I,k,n}}{\hat{D}_{k,n}}$

where

- k is the threshold parameter,

- n is the sample size,

- $X_1, ..., X_n$ describes the `sample`, each X_s is a d-dimensional vector $X_s^{(t)}$ for $t = 1, ..., d$,

- $R_s^{(t)}$ denotes the rank of $X_s^{(t)}$ among $X_1^{(t)}, ..., X_n^{(t)}$,

- and $\overline{R}_s^{(t)} = \min((n - R_s^{(t)} + 1)/k, 1)$.

Value

The function returns a list of two elements:

- `subsets` A list of subsets from $\{1, ..., d\}$.

  When `ind` is given as an integer, `subsets` is the list of subsets from $\{1, ..., d\}$ with cardinality `ind`. When `ind` is the list, it corresponds to `subsets`.

  When `ind = "with.singletons"` `subsets` is the list of all non empty subsets in $\{1, ..., d\}$.

  When `ind = "all"` `subsets` is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- `tic` A vector of tail importance coefficients, or their sobol versions when `sobol = "TRUE"`.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

[tic](#) and [tsicEmp](#)

Examples

```
## Fix a 5-dimensional asymmetric tail dependence structure
(ds5 <- gen.ds(d = 5))

## Generate a 1000-sample of Mevlog random vectors associated with ds5
sample5 <- rMevlog(n = 1000, ds = ds5)

## Compute empirical tic values according cardinality
res2 <- ticEmp(sample5, ind = 2, k = 100, sobol = TRUE)
res3 <- ticEmp(sample5, ind = 3, k = 100, sobol = TRUE)
res4 <- ticEmp(sample5, ind = 4, k = 100, sobol = TRUE)

## Represent the empirical indices associated with pairs
barplot(res2$tic ~ as.character(res2$subsets), las = 2,
     xlab = "", ylab = "", main = "Tail Sobol Indices (cardinality 2)")

## Represent the empirical indices associated with triplets
barplot(res3$tic ~ as.character(res3$subsets), las = 2,
     xlab = "", ylab = "", main = "Tail Sobol Indices (cardinality 3)")

## Represent the empirical indices associated with quadriplets
barplot(res4$tic ~ as.character(res4$subsets), las = 2,
     xlab = "", ylab ="", main = "Tail Sobol Indices (cardinality 4)")

## Check the sum-to-one constraint of empirical tail Sobol indices
sum(ticEmp(sample5, ind = "with.singletons", k = 100,  sobol = TRUE)$tic)
```

---

tsic	*Tail superset importance coefficients for Mevlog models.*

Description

Tail superset importance coefficients for Mevlog models. A `Mevlog` model is a multivariate extreme value (symmetric or asymmetric) logistic model.

Usage

```
tsic(ds, ind = 2, n.MC = 1000, sobol = FALSE, norm = FALSE)
```

Arguments

ds	An object of class ds.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = 2, all pairwise coefficients are computed.
n.MC	Monte Carlo sample size. Default value is 1000. See Details.

sobol   A boolean. 'FALSE' (the default). If 'TRUE': the index is normalized by the
               theoretical global variance.

norm   A boolean. 'FALSE' (the default): original tsic is computed. 'TRUE': tsic is
               normalized by its upper bound.

Details

The tail dependence structure is specified using a ds object, which corresponds to the stable tail dependence function ℓ. The process for deducing the stable tail dependence function ℓ from ds is explained in the Details section of gen.ds.

A tail superset importance coefficient (tsic) is a measure of the importance of a subset of components (and their supersets) in contributing to the global variance decomposition of ℓ. The tsic is computed using Monte Carlo methods based on the integral formula (3) in Mercadier and Roustant (2019). Recall that Formula (9) in Liu and Owen (2006) provides an integral representation of the superset importance coefficient.

The tail dependograph is plotted using pairwise tsic values, which are computed using the function tsic and the ind = 2 option.

The upper bound for a tsic associated with subset I is given by Theorem 2 in Mercadier and Ressel (2021). If $|I|$ is the cardinality of subset I, then the upper bound is $2(|I|!)^2/((2|I| + 2)!)$.

The tail dependence structure is set by a ds object. It thus corresponds to the stable tail dependence function ℓ.

Value

The function returns a list of two elements

- subsets A list of subsets from $\{1, ..., d\}$.

  When ind is given as an integer, subsets is the list of subsets from $\{1, ..., d\}$ with cardinality ind.

  When ind is a list, it corresponds to subsets.

  When ind = "with.singletons" subsets is the list of all non empty subsets in $\{1, ..., d\}$.

  When ind = "all" subsets is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- tsic A vector of tail superset importance coefficients associated with the list subsets. When norm = TRUE, then tsic are normalized in the sense that the original values are divided by corresponding upper bounds.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Liu, R. and Owen, A. B. (2006) Estimating mean dimensionality of analysis of variance decompositions. *J. Amer. Statist. Assoc.*, **101(474)**:712–721.

Mercadier, C. and Ressel, P. (2021) Hoeffding–Sobol decomposition of homogeneous co-survival functions: from Choquet representation to extreme value theory application. Dependence Modeling, **9(1)**, 179–198.

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

Smith, R. L. (1990) Max-stable processes and spatial extremes. *Dept. of Math., Univ. of Surrey*, Guildford GU2 5XH, England.

Tiago de Oliveira, J. (1962/63) Structure theory of bivariate extremes, extensions. *Estudos de Matematica, Estatistica, e Economicos*, 7:165–195.

See Also

graphs, ellMevlog

Examples

```
## Fix a 5-dimensional asymmetric tail dependence structure
ds5 <- gen.ds(d = 5)

## Compute pairwise tsic
tsic(ds = ds5, ind = 2)

## Plot the tail dependograph
graphs(ds = ds5)

## Compute tsic on two specific subsets
tsic(ds = ds5, ind = list(1:4, 3:5))

## Compute normalized version of tsic
tsic(ds5,  ind = list(1:4, 3:5), norm = TRUE)

## Compute Sobol and normalized version of tsic
tsic(ds5,  ind = list(1:4, 3:5), norm = TRUE, sobol = TRUE)
```

---

tsicEmp   *Empirical tail superset importance coefficients.*

Description

Computes on a sample the tail superset importance coefficients (tsic) associated with threshold k. The value may be renormalized by the empirical global variance (Sobol version) and/or by its theoretical upper bound.

Usage

```
tsicEmp(sample, ind = 2, k, sobol = FALSE, norm = FALSE)
```

Arguments

sample	A (n times d) matrix.
ind	A character string among "with.singletons" and "all" (without singletons), or an integer in $\{2, ..., d\}$ or a list of subsets from $\{1, ..., d\}$. The default is ind = 2, all pairwise coefficients are computed.
k	An integer smaller or equal to n.
sobol	A boolean. 'FALSE' (the default). If 'TRUE': the index is normalized by the empirical global variance.
norm	A boolean. 'FALSE' (the default). If 'TRUE': the index is normalized by its theoretical upper bound.

Details

The theoretical functional decomposition of the variance of the stdf ℓ consists in writing $D(\ell) = \sum_{I \subseteq \{1,...,d\}} D_I(\ell)$ where $D_I(\ell)$ measures the variance of $\ell_I(U_I)$ the term associated with subset I in the Hoeffding-Sobol decomposition of ℓ ; note that U_I represents a random vector with independent standard uniform entries.

Fixing a subset of components I, the theoretical tail superset importance coefficient is defined by $\Upsilon_I(\ell) = \sum_{J \supseteq I} D_J(\ell)$. A theoretical upper bound for tsic $\Upsilon_I(\ell)$ is given by Theorem 2 in Mercadier and Ressel (2021) which states that $\Upsilon_I(\ell) \leq 2(|I|!)^2/((2|I|+2)!)$.

Here, the function tsicEmp evaluates, on a n-sample and threshold k, the empirical tail superset importance coefficient $\hat{\Upsilon}_{I,k,n}$ the empirical counterpart of $\Upsilon_I(\ell)$.

Under the option sobol = TRUE, the function tsicEmp returns $\dfrac{\hat{\Upsilon}_{I,k,n}}{\hat{D}_{k,n}}$ the empirical counterpart of $\dfrac{\Upsilon_I(\ell)}{D_I(\ell)}$.

Under the option norm = TRUE, the quantities are multiplied by $\dfrac{(2|I|+2)!}{2(|I|!)^2}$.

Proposition 1 and Theorem 2 of Mercadier and Roustant (2019) provide several rank-based expressions

$$\hat{\Upsilon}_{I,k,n} = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t \in I} (\min(\overline{R}_s^{(t)}, \overline{R}_{s'}^{(t)}) - \overline{R}_s^{(t)} \overline{R}_{s'}^{(t)}) \prod_{t \notin I} \min(\overline{R}_s^{(t)}, \overline{R}_{s'}^{(t)})$$

$$\hat{D}_{k,n} = \frac{1}{k^2} \sum_{s=1}^n \sum_{s'=1}^n \prod_{t \in I} \min(\overline{R}_s^{(t)}, \overline{R}_{s'}^{(t)}) - \prod_{t \in I} \overline{R}_s^{(t)} \overline{R}_{s'}^{(t)}$$

where

- k is the threshold parameter,
- n is the sample size,
- $X_1, ..., X_n$ describes the sample, each X_s is a d-dimensional vector $X_s^{(t)}$ for $t = 1, ..., d$,
- $R_s^{(t)}$ denotes the rank of $X_s^{(t)}$ among $X_1^{(t)}, ..., X_n^{(t)}$,
- and $\overline{R}_s^{(t)} = \min((n - R_s^{(t)} + 1)/k, 1)$.

Value

The function returns a list of two elements:

- subsets A list of subsets from $\{1, ..., d\}$.

  When ind is given as an integer, subsets is the list of subsets from $\{1, ..., d\}$ with cardinality ind. When ind is the list, it corresponds to subsets.

  When ind = "with.singletons" subsets is the list of all non empty subsets in $\{1, ..., d\}$.

  When ind = "all" subsets is the list of all subsets in $\{1, ..., d\}$ with cardinality larger or equal to 2.

- tsic A vector of empirical tail superset importance coefficients associated with the list subsets. When norm = TRUE, then tsic are normalized in the sense that the original values are divided by corresponding upper bounds.

Author(s)

Cécile Mercadier (mercadier@math.univ-lyon1.fr)

References

Mercadier, C. and Ressel, P. (2021) Hoeffding–Sobol decomposition of homogeneous co-survival functions: from Choquet representation to extreme value theory application. Dependence Modeling, **9(1)**, 179–198.

Mercadier, C. and Roustant, O. (2019) The tail dependograph. *Extremes*, **22**, 343–372.

See Also

graphsEmp, ellEmp

Examples

```
## Fix a 6-dimensional asymmetric tail dependence structure
ds <- gen.ds(d = 6, sub = list(1:4,5:6))

## Plot the  tail dependograph
graphs(ds)

## Generate a 1000-sample of Archimax Mevlog random vectors
## associated with ds and underlying distribution exp
sample <- rArchimaxMevlog(n = 1000, ds = ds, dist = "exp", dist.param = 1.3)

## Compute tsic values associated with subsets
## of cardinality 2 or more \code{ind = "all"}
res <- tsicEmp(sample = sample, ind = "all", k = 100, sobol = TRUE, norm = TRUE)

## Select the significative tsic
indices_nonzero <- which(res$tsic %in% boxplot.stats(res$tsic)$out == TRUE)

## Subsets associated with significative tsic reflecting the tail support
as.character(res$subsets[indices_nonzero])
```

```
## Pairwise tsic are obtained by
res_pairs <- tsicEmp(sample = sample, ind = 2, k = 100, sobol = TRUE, norm = TRUE)

## and plotted in the tail dependograph
graphsEmp(sample, k = 100)
```

Index