Package ‘flood’

October 13, 2022

Type Package

Title Statistical Methods for the (Regional) Analysis of Flood Frequency

Version 0.1.1

Description Includes several statistical methods for the estimation of parameters and high quantiles of river flow distributions. The focus is on regional estimation based on homogeneity assumptions and computed from multivariate observations (multiple measurement stations). For details see Kinsvater et al. (2017) <arXiv:1701.06455>.

Imports evd, TLMoments, magrittr, copula

License GPL (>= 2)

RoxygenNote 5.0.1

NeedsCompilation no

Author Friederike Deiters [aut, cre], Paul Kinsvater [aut]

Maintainer Friederike Deiters <deiters@statistik.tu-dortmund.de>

Repository CRAN

Date/Publication 2017-01-30 14:50:01

R topics documented:

`confInt_sTL` *Estimated confidence intervals using sTL*

Description

Estimated regional (or local) (1-alpha)-confidence intervals for an estimated quantile by using seasonal TL(0,1)-moments (sTL).

Usage

```
confInt_sTL(x1, x2, p, j = 1, alpha = 0.05)
```

Arguments

`x1`	vector or matrix of observations from season 1 (rows: observations, columns: stations).
`x2`	vector or matrix of observations from season 2 (rows: observations, columns: stations).
`p`	a probability.
`j`	quantile and parameter estimation for the jth station (jth column of `x`). Irrelevant if is `x1` and `x2` are vectors.
`alpha`	confidence level for confidence interval.

Value

List of

- `ci` confidence interval.
- `quant` estimated quantile from a two-component GEV using trimmed L-moments (leftrim=0, rightrim=1).
- `var` variance of the estimated quantile.

Examples

```
library("evd")
# Seasonal observations of 80 years at one station:
x1 <- rgev(80, 2, 1, 0.2) # observations from season 1
x2 <- rgev(80, 3, 1, 0.3) # observations from season 2
confInt_sTL(x1=x1, x2=x2, p=0.95, alpha=0.05)

# Seasonal observations of 100 years at 4 stations:
x1 <- matrix(rgev(400, 2, 1, 0.3), ncol=4)
x2 <- matrix(rgev(400, 4, 1, 0.2), ncol=4)
confInt_sTL(x1=x1, x2=x2, j=2, p=0.95, alpha=0.05)
```

confInt_TL — *Estimated confidence intervals using TL*

Description

Estimated regional (or local) (1-alpha)-confidence intervals for an estimated quantile by using annual TL(0,1)-moments (TL).

Usage

```
confInt_TL(x, p, j = 1, alpha = 0.05)
```

Arguments

`x`	vector or matrix of annual observations.
`p`	a probability.
`j`	quantile and parameter estimation for the jth station (jth column of `x`). Irrelevant if is `x` is a vector.
`alpha`	confidence level for confidence interval.

Value

List of

- `ci` confidence interval.
- `quant` estimated quantile from a GEV using trimmed L-moments ($leftrim = 0, rightrim = 1$).

Examples

```
library("evd")
# Seasonal observations of 80 years at one station:
x1 <- rgev(80, 2, 1, 0.2) # observations from season 1
x2 <- rgev(80, 3, 1, 0.3) # observations from season 2
x <- seas2ann(x1, x2) # calculaes annual maxima of the two seasons
confInt_TL(x=x, p=0.95, alpha=0.05)

# Seasonal observations of 100 years at 4 stations:
x1 <- matrix(rgev(400, 2, 1, 0.3), ncol=4) # observations from season 1
x2 <- matrix(rgev(400, 2, 1, 0.2), ncol=4) # observations from season 2
x <- seas2ann(x1, x2) # calculaes annual maxima of the two seasons
confInt_TL(x=x, j=2, p=0.95, alpha=0.05)
```

confInt_W *Estimated confidence intervals using W*

Description

Estimated regional (or local) (1-alpha)-confidence intervals for an estimated quantile by using the annual Weissman estimator (W).

Usage

```
confInt_W(x, p, j = 1, alpha = 0.05, ...)
```

Arguments

x	vector or matrix of annual observations.
p	a probability. Should be between $1-k_j/n_j$ and 1, where n_j is the sample length of the j-th column.
j	quantile and parameter estimation for the jth station (jth column of x). Irrelevant if is x1 and x2 are vectors.
alpha	confidence level for confidence interval.
...	additional arguments, see RegioWeissman

Value

List of

- ci confidence interval.
- quant estimated quantile using Weissman's estimator.

Examples

```
library("evd")
# Seasonal observations of 80 years at one station:
x1 <- rgev(80, 2, 1, 0.2) # observations from season 1
x2 <- rgev(80, 3, 1, 0.3) # observations from season 2
x <- seas2ann(x1, x2) # calculaes annual maxima of the two seasons
confInt_W(x=x, p=0.95, alpha=0.05)

# Seasonal observations of 100 years at 4 stations:
x1 <- matrix(rgev(400, 2, 1, 0.3), ncol=4) # observations from season 1
x2 <- matrix(rgev(400, 2, 1, 0.2), ncol=4) # observations from season 2
x <- seas2ann(x1, x2) # calculaes annual maxima of the two seasons
confInt_W(x=x, j=2, p=0.95, alpha=0.05)
```

dGEVxGEV *Two-component generalized extreme value distribution (GEV)*

Description

Density, distribution function, quantile function and random generation for a two-component GEV distribution (product of two GEVs).

Usage

```
dGEVxGEV(x, param1, param2)

pGEVxGEV(q, param1, param2)

qGEVxGEV(p, param1, param2)

rGEVxGEV(n, param1, param2)
```

Arguments

x	vector of quantiles.
param1	three-dimensional vector (loc, scale, shape)' of a GEV from season 1.
param2	three-dimensional vector (loc, scale, shape)' of a GEV from season 2.
q	vector of quantiles.
p	vector of probabilities.
n	number of observations.

Details

These functions use the parametrization of the gev-functions from the package 'evd'. The distribution F of a two-component GEV is: $F = F_1 \cdot F_2$, where F_1 and F_2 are two distribution functions of a GEV.

Examples

```
# density and distribution function of a two-component GEV:
par(mfrow=c(3,1))
curve(dGEVxGEV(x, c(2,1,0.2), c(3,2,0.4)), from=0, to=20, ylab="Density", n=1000)
curve(pGEVxGEV(x, c(2,1,0.2), c(3,2,0.4)), from=0, to=20, ylab="Probability", n=1000)

# quantiles of a two-component GEV:
qGEVxGEV(p=c(0.9, 0.99), c(2,1,0.2), c(3,2,0.4))

# random numbers of a two-component GEV:
set.seed(23764)
rn <- rGEVxGEV(1000, c(2,1,0.1), c(3,2,0))
hist(rn, breaks=40, freq=FALSE, main="")
```

```
curve(dGEVxGEV(x, c(2,1,0.1), c(3,2,0)), from=0, to=20,
ylab="density", n=1000, col="red", add=TRUE)
```

hill *Hill's estimator*

Description

Estimation of heavy tails with Hill's estimator

Usage

```
hill(x, k)
```

Arguments

x	Vector or matrix of observations
k	Number of relative excesses involved in the estimation of the extreme value index gamma. If k is missing, it will be set to • $k = \lfloor 2n^{2/3} \rfloor$, where n is the sample length of the vector x after removing missing values • $k = \left\lfloor \frac{2n^{2/3}}{d^{1/3}} \right\rfloor$, where d is the number of columns of the matrix x and n the length of each column after removing missing values.

Value

Hill's estimator for each sample.

Examples

```
library("evd")
x1 <- rgev(100, loc = 2, scale = 1, shape=0.4)
hill(x1, k=20)
x2 <- rgev(100, loc = 2, scale = 1, shape=0.5)
hill(cbind(x1, x2), k = c(20, 25))
x2[c(4,8,39)] <- NA
hill(cbind(x1, x2), k=c(20, 25))
# if leaving out k, it will be set to floor(2*n^(2/3)/d^(1/3)) = c(34,33):
hill(cbind(x1, x2)) # is the same as:
hill(cbind(x1, x2), k=c(34,33))
```

qBM — *Block maxima distribution*

Description

Calculates quantiles of a block maxima distribution.

Usage

```
qBM(p, b, param)
```

Arguments

p	vector of probabilities.
b	block length (in general b ≥ 2).
param	three-dimensional vector with location (mu), scale (sigma) and shape (xi) parameter.

Details

Formular of a block maxima distribution function:

$$F_j(x) = F_j^{(b)}(x) = \left[2 \cdot T_{1/\xi} \left(\left\{ 1 + \xi \frac{x - \mu_j}{\sigma_j} \right\} \cdot T_{1/\xi}^{-1} \left(1 - \frac{1}{2b} \right) \right) - 1 \right]^b ,$$

where T_ν denote the t-distribution function with ν degrees of freedom.

Value

Quantile of a block maxima distribution.

Examples

```
qBM(p=c(0.75, 0.99), b=12, param=c(2,1,0.2))
```

RegioGEV — *Regional (or local) parameter and quantile estimation*

Description

Calculates regional (or local) parameters of a generalized extreme value (GEV) distribution using (trimmed) L-moments (see TLMoments and parameters) from a vector or matrix of observation. Based on these parameters, a p-quantile of the GEV will be calculated for the jth station.

Usage

```
RegioGEV(x, p, j = 1, leftrim = 0, rightrim = 0, na.rm = TRUE, ...)
```

Arguments

x	vector or matrix of observations (rows: observations, d columns: stations).
p	a probability.
j	quantile and parameter estimation for the jth station (jth column of x). Irrelevant if is x is a vector.
leftrim	integer indicating lower trimming parameter (≥ 0).
rightrim	integer indicating upper trimming parameter (≥ 0).
na.rm	Should missing values be removed?
...	additional arguments, see TLMoments.

Details

The optimal weights will be calculated as described in "Kinsvater, Fried and Lilienthal (2015): Regional extreme value index estimation and a test of tail homogeneity, Environmetrics, DOI: 10.1002/env.2376, Section 3.2". If it's not possible to calculate optimal weights (negative eigenvaules of an estimated covarinace matrix), simple weights will be calculated: $w_j = \frac{n_j}{sum_{j=1}^d n_j}$

Value

List of

- quant quantile calculation from an estimated GEV with a regional shape-parameter.
- param estimated parameter vector from a GEV (using L-moments or trimmed L-moments).
- w optimal or simple weighting (just returned if x is a matrix).

Examples

```
library("evd")
# sample observations of 75 years at one station:
x <- rgev(75) # x is a vector
RegioGEV(x=x, p=0.95)

x2 <- c(NA, NA, x[1:60], NA, x[61:75]) # vector of observations with missing values
RegioGEV(x=x2, p=0.95) # NAs will be removed

# sample observations of 100 years at 4 stations:
set.seed(1053)
x <- matrix(rgev(400, 2, 1, 0.3), ncol=4)
RegioGEV(x=x, p=0.9, j=3, leftrim=0, rightrim=0) # optimal weighting
RegioGEV(x=x, p=0.9, j=3, leftrim=0, rightrim=1) # optimal weighting

# With missing values:
x2 <- x
x2[c(54, 89, 300)] <- NA
RegioGEV(x=x2, p=0.9, j=3, leftrim=0, rightrim=0)

# sample again observations of 100 years at 4 stations:
set.seed(958)
```

```
x <- matrix(rgev(400, 2, 1, 0.3), ncol=4)
RegioGEV(x=x, p=0.9, j=3, leftrim=0, rightrim=0) # simple weighting
```

RegioGEVSeas *Seasonal regional (or local) parameter and quantile estimation*

Description

Calculates regional (or local) parameters of a two-component GEV distribution (product of two GEVs) using (trimmed) L-moments (see TLMoments and parameters) from two vectors or two matrices of observation, e.g. winter and summer observations from one or more than one station. Based on these two parameter vectors, a p-quantile of the two-component GEV will be calculated for the jth station.

Usage

```
RegioGEVSeas(x1, x2, p, j = 1, leftrim = 0, rightrim = 0, na.rm = TRUE,
  ...)
```

Arguments

`x1`	vector or matrix of observations from season 1 (rows: observations, columns: stations).
`x2`	vector or matrix of observations from season 2 (rows: observations, columns: stations).
`p`	a probability.
`j`	quantile and parameter estimation for the jth station (jth column of `x`). Irrelevant if is `x1` and `x2` are vectors.
`leftrim`	integer indicating lower trimming parameter (≥ 0).
`rightrim`	integer indicating upper trimming parameter (≥ 0).
`na.rm`	Should missing values be removed?
`...`	additional arguments, see TLMoments.

Value

List of

- `quant` quantile calculation from an estimated two-component GEV with a regional (or local) shape-parameters.
- `param1` estimated parameter vector from season 1 from a GEV (using L-moments or trimmed L-moments).
- `param2` estimated parameter vector from season 2 from a GEV (using L-moments or trimmed L-moments).

Examples

```
library("evd")
# Seasonal observations of 80 years at one station:
x1 <- rgev(80, 2, 1, 0.2) # observations from season 1
x2 <- rgev(80, 3, 1, 0.3) # observations from season 2
RegioGEVSeas(x1=x1, x2=x2, p=0.95)

# Missing values in both seasons in the same year(s):
x1a <- c(NA, x1, NA)
x2a <- c(NA, x2, NA)
RegioGEVSeas(x1a, x2a, p=0.99, leftrim=0, rightrim=0, na.rm=TRUE)

# Missing values in both seasons in different year(s):
x1b <- x1
x1b[c(4,19)] <- NA
x2b <- x2
x2b[77] <- NA
RegioGEVSeas(x1b, x2b, p=0.99, leftrim=0, rightrim=0, na.rm=TRUE)

# Seasonal observations of 100 years at 4 stations:
x1 <- matrix(rgev(400, 2, 1, 0.3), ncol=4)
x2 <- matrix(rgev(400, 4, 1, 0.2), ncol=4)
# estimate quantile for station 1 and 2 (consider the same shape-parameters):
RegioGEVSeas(x1, x2, p=0.99, j=1, leftrim=0, rightrim=0)
RegioGEVSeas(x1, x2, p=0.99, j=2, leftrim=0, rightrim=0)

# With missing values:
x3 <- x1
x4 <- x2
x3[c(54, 89, 300)] <- NA
RegioGEVSeas(x3, x4, p=0.99, j=1, leftrim=0, rightrim=0)
```

RegioHill *Regional EVI estimator*

Description

Estimation of the positive extreme value index (EVI) based on multiple local Hill estimators. We assume heavy-tail homogeneity, i.e., all local EVI's are the same.

Usage

```
RegioHill(x, k, k.qu = 20, type = "evopt", alpha = 0.05, ci = "nonlog")
```

Arguments

x	Vector or matrix of observations
k	Number of relative excesses involved in the estimation of the extreme value index gamma. If k is missing, it will be set to

- $k = \lfloor 2n^{2/3} \rfloor$, where n is the sample length of the vector x after removing missing values
- $k = \left\lfloor \frac{2n^{2/3}}{d^{1/3}} \right\rfloor$, where d is the number of columns of the matrix x and n the length of each column after removing missing values.

`k.qu` Tuning parameter for estimation of empirical variance; only needed if `type="opt"`.

`type` Choose either `"evopt"` if extreme value dependent, `"ind"` if independent or `"opt"` for arbitrarily dependent components.

`alpha` Confidence level for confidence interval.

`ci` Either `"nonlog"` for standart or `"log"` for non-standart confidence interval based on log-transformed hill estimates.

Value

List of

- `est` a weighted average of local Hill estimates.
- `Sigma` an estimate of the corresponding variance matrix.
- `CI` a confidence interval.

Examples

```
library("evd")
x1 <- rgev(150, loc = 2, scale = 1, shape=0.4)
hill(x1, k=20)
x2 <- rgev(100, loc = 2.5, scale = 1, shape=0.4)
x2 <- c(x2, rep(NA, 50))
x <- cbind(x1, x2)
k <- c(40, 30)
RegioHill(x, k)
```

RegioWeissman *Quantile estimation: Weissman's extrapolation*

Description

Estimation of the p-quantile based on multiple local Hill estimators and Weissman's extrapolation formula. We assume heavy-tail homogeneity, i.e., all local EVI's are the same.

Usage

```
RegioWeissman(x, j = 1, p, k, k.qu = 20, type = "evopt", alpha = 0.05)
```

Arguments

- `x` — Vector or matrix of observations
- `j` — The number of the target site, i.e., if `j=2` the p-quantile of the second column of `x` is estimated.
- `p` — The probability of interest; should be between $1 - k_j/n_j$ and 1, where n_j is the sample length of the j-th column.
- `k` — Number of relative excesses involved in the estimation of the extreme value index gamma. If `k` is missing, it will be set to
  - $k = \lfloor 2n^{2/3} \rfloor$, where n is the sample length of the vector `x` after removing missing values
  - $k = \left\lfloor \frac{2n^{2/3}}{d^{1/3}} \right\rfloor$, where d is the number of columns of the matrix `x` and n the length of each column after removing missing values.
- `k.qu` — Tuning parameter for estimation of empirical variance; only needed if `type="opt"`.
- `type` — Choose either `"evopt"` if extreme value dependent, `"ind"` if independent or `"opt"` for arbitrarily dependent components.
- `alpha` — Confidence level for confidence interval.

Value

List of

- `est` Point estimate of p-quantile of column j
- `CI` the corresponding alpha-confidence interval
- `EVI` estimate of the extreme value index
- `k` tail sample size
- `p` a probability
- `u.kn` (n-k)-th largest observation, where n is the sample length at station j after removing missing values
- `description` a short description.

Examples

```
library("evd")
# sample observations of 75 years at one station:
x <- rgev(75, 0, 1, 0) # x is a vector
RegioWeissman(x=x, p=0.95)

x2 <- c(NA, NA, x[1:60], NA, x[61:75]) # vector of observations with missing values
RegioWeissman(x=x2, p=0.95) # NAs will be removed

# sample observations of 100 years at 4 stations:
set.seed(1053)
x <- matrix(rgev(400, 2, 1, 0.3), ncol=4)
RegioWeissman(x=x, p=0.9, j=3)
```

```
# With missing values:
x2 <- x
x2[sample(100, 12),2] <- NA
RegioWeissman(x=x2, p=0.9, j=3)
```

`RegioWeissmanSeas` *Quantile estimation: Weissman's extrapolation for seasonal data*

Description

Estimation of the p-quantile based on multiple local Hill estimators and Weissman's extrapolation formula with seasonality. We assume heavy-tail homogeneity within each season, i.e., all local EVI's are the same.

Usage

```
RegioWeissmanSeas(x, j = 1, p, ...)
```

Arguments

`x`	List of 2 elements: each element is consistent a vector or matrix of observations
`j`	The number of the target site, i.e., if `j=2` the p-quantile of the second station is estimated.
`p`	The probability of interest; should be between $1 - k_j/n_j$ and 1, where n_j is the sample length of the `j`-th column.
`...`	additional arguments, see RegioWeissman

Value

Point estimate of seasonal p-quantile of column `j` and a short description.

Examples

```
library("evd")
# Local & seasonal (observations of 80 years at one station):
x1 <- rgev(80, 2, 1, 0.2) # observations from season 1
x2 <- rgev(80, 3, 1, 0.3) # observations from season 2
x <- list(x1, x2)
RegioWeissmanSeas(x=x, j=1, p=0.99)

x1 <- matrix(rgev(400, 3, 1, 0.1), ncol=4) # observations of season 1 at 4 stations
x2 <- matrix(rgev(400, 2, 1, 0.3), ncol=4) # observations of season 2 at 4 stations
x <- list(x1, x2)
RegioWeissmanSeas(x=x, j=1, p=0.99)
```

seas2ann *Annual maxima from seasonal maxima*

Description

Calculates annual maxima from seasonal maxima of two seasons.

Usage

```
seas2ann(x1, x2, na.rm = TRUE)
```

Arguments

x1	vector or matrix of observations from season 1 (rows: observations, columns: stations).
x2	vector or matrix of observations from season 2 (rows: observations, columns: stations).
na.rm	logical. TRUE: the annual maximum will be calculated even if one observation of the two seasons is a missing value, e.g. winter maximum is 58 and summer maximum is NA the annual maximum is 58. If both observations are NA, the annual maximum is set to NA, too. FALSE: the annual maximum will be set to NA if one observation of the two seasons is a missing value, e.g. winter maximum is 58 and summer maximum is NA the annual maximum is NA.

Value

Matrix of annual observations (rows: observations, columns: stations).

Examples

```
set.seed(28379)
x1 <- matrix(round(rnorm(8, 20, 25)), ncol=2)
x1[2] <- NA
x2 <- matrix(round(rnorm(8, 20, 25)), ncol=2)
x2[c(2,5,6)] <- NA
x1
x2
seas2ann(x1,x2,TRUE)
seas2ann(x1,x2,FALSE)
```

TailAnova *Heavy-tail ANOVA*

Description

A test of heavy-tail homogeneity, that is, equality of the positive extreme value index for all d columns of x.

Usage

```
TailAnova(x, k, k.qu = 20, type = "evopt", cf = TRUE)
```

Arguments

x	Matrix of observations
k	Number of relative excesses involved in the estimation of the extreme value index gamma. If k is missing, it will be set to $k = \left\lfloor \frac{2n^{2/3}}{d^{1/3}} \right\rfloor$, where d is the number of columns of the matrix x and n the length of each column after removing missing values.
k.qu	Tuning parameter for estimation of empirical variance; only needed if type="opt".
type	Choose either "evopt" if extreme value dependent, "ind" if independent or "opt" for arbitrarily dependent components.
cf	If TRUE, a correctur factor is used, which improves the size at the cost of power.

Value

Test statistic and p-value.

Examples

```
library("evd")
set.seed(6754)
x1 <- rgev(150, loc = 2, scale = 1, shape=0.4)
x2 <- rgev(150, loc = 2.5, scale = 1, shape=0.1) # H_0 violated because of different shapes
x <- cbind(x1, x2)
TailAnova(x)

x1 <- rgev(150, loc = 2, scale = 1, shape=0.3)
x2 <- rgev(150, loc = 2.5, scale = 1, shape=0.3) # H_0 not violated because of same shapes
x <- cbind(x1, x2)
TailAnova(x)
```

`xiAnova` *Homogeneity test for the shape*

Description

A test for assumption H_0: "shape parameter is equal for all d GEV margins" with test statistic based on (trimmed) L-moments.

Usage

```
xiAnova(x, leftrim = 0, rightrim = 1)
```

Arguments

`x`	matrix of observations (rows: observations, d columns: stations).
`leftrim`	integer indicating lower trimming parameter (≥ 0).
`rightrim`	integer indicating upper trimming parameter (≥ 0).

Value

p-value of the test.

Examples

```
library("evd")
# sample observations of 100 years at 5 stations:
set.seed(1053)
x19 <- matrix(rgev(400, 2, 1, 0.1), ncol=4) # 4 stations with the same shape
x10 <- rgev(100, 2, 1, 0.4) # one station with a different shape
x <- cbind(x19, x10)
xiAnova(x=x, leftrim=0, rightrim=1)
```

Index