Package 'homtest'

February 20, 2015

Version 1.0-5

Date 2009-03-26

Title Homogeneity tests for Regional Frequency Analysis

Author Alberto Viglione

Maintainer Alberto Viglione <viglione@hydro.tuwien.ac.at>

Description A collection of homogeneity tests described in: Viglione
A., Laio F., Claps P. (2007) ``A comparison of homogeneity
tests for regional frequency analysis'', Water Resources
Research, 43, W03428, doi:10.1029/2006WR005095. More on
Regional Frequency Analysis can be found in package nsRFA.

Depends stats

License GPL (>= 2)

URL http://www.idrologia.polito.it/~alviglio

Repository CRAN

Date/Publication 2012-11-05 10:30:42

NeedsCompilation no

R topics documented:

---

annualflows   *Data-sample*

Description

Total annual flow, expressed in mm, of 47 stations in Piemonte (Italy).

Usage

```
 annualflows
```

Format

Data.frame containing annual flow data of 47 stations.

Examples

```
data(annualflows)
annualflows
summary(annualflows)
x <- annualflows["dato"][,]
cod <- annualflows["cod"][,]
split(x,cod)
sapply(split(x,cod),mean)
sapply(split(x,cod),median)
sapply(split(x,cod),quantile)
sapply(split(x,cod),Lmoments)
```

GEV   *Three parameter generalized extreme value distribution and L-moments*

Description

GEV provides the link between L-moments of a sample and the three parameter generalized extreme value distribution.

Usage

```
f.GEV (x, xi, alfa, k)
F.GEV (x, xi, alfa, k)
invF.GEV (F, xi, alfa, k)
Lmom.GEV (xi, alfa, k)
par.GEV (lambda1, lambda2, tau3)
rand.GEV (numerosita, xi, alfa, k)
```

Arguments

x	vector of quantiles
xi	vector of GEV location parameters
alfa	vector of GEV scale parameters
k	vector of GEV shape parameters
F	vector of probabilities
lambda1	vector of sample means
lambda2	vector of L-variances
tau3	vector of L-CA (or L-skewness)
numerosita	numeric value indicating the length of the vector to be generated

Details

See http://en.wikipedia.org/wiki/Generalized_extreme_value_distribution for an introduction to the GEV distribution.

Definition

Parameters (3): ξ (location), α (scale), k (shape).

Range of x: $-\infty < x \le \xi + \alpha/k$ if $k > 0$; $-\infty < x < \infty$ if $k = 0$; $\xi + \alpha/k \le x < \infty$ if $k < 0$.

Probability density function:

$$f(x) = \alpha^{-1} e^{-(1-k)y - e^{-y}}$$

where $y = -k^{-1} \log\{1 - k(x - \xi)/\alpha\}$ if $k \ne 0$, $y = (x - \xi)/\alpha$ if $k = 0$.

Cumulative distribution function:

$$F(x) = e^{-e^{-y}}$$

Quantile function: $x(F) = \xi + \alpha[1 - (-\log F)^k]/k$ if $k \ne 0$, $x(F) = \xi - \alpha \log(-\log F)$ if $k = 0$.

$k = 0$ is the Gumbel distribution; $k = 1$ is the reverse exponential distribution.

L-moments

L-moments are defined for $k > -1$.

$$\lambda_1 = \xi + \alpha[1 - \Gamma(1 + k)]/k$$

$$\lambda_2 = \alpha(1 - 2^{-k})\Gamma(1 + k)]/k$$

$$\tau_3 = 2(1 - 3^{-k})/(1 - 2^{-k}) - 3$$

$$\tau_4 = [5(1 - 4^{-k}) - 10(1 - 3^{-k}) + 6(1 - 2^{-k})]/(1 - 2^{-k})$$

Here Γ denote the gamma function

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

Parameters

To estimate k, no explicit solution is possible, but the following approximation has accurancy better than 9×10^{-4} for $-0.5 \leq \tau_3 \leq 0.5$:

$$k \approx 7.8590c + 2.9554c^2$$

where

$$c = \frac{2}{3 + \tau_3} - \frac{\log 2}{\log 3}$$

The other parameters are then given by

$$\alpha = \frac{\lambda_2 k}{(1 - 2^{-k})\Gamma(1 + k)}$$

$$\xi = \lambda_1 - \alpha[1 - \Gamma(1 + k)]/k$$

Value

f.GEV gives the density f, F.GEV gives the distribution function F, invF.GEV gives the quantile function x, Lmom.GEV gives the L-moments (λ_1, λ_2, τ_3, τ_4), par.GEV gives the parameters (xi, alfa, k), and rand.GEV generates random deviates.

Note

Lmom.GEV and par.GEV accept input as vectors of equal length. In f.GEV, F.GEV, invF.GEV and rand.GEV parameters (xi, alfa, k) must be atomic.

Author(s)

Alberto Viglione, e-mail: <alviglio@tiscali.it>.

References

Hosking, J.R.M. and Wallis, J.R. (1997) Regional Frequency Analysis: an approach based on L-moments, Cambridge University Press, Cambridge, UK.

See Also

rnorm, runif, KAPPA, Lmoments.

GUMBEL   *Two parameter Gumbel distribution and L-moments*

Description

GUMBEL provides the link between L-moments of a sample and the two parameter Gumbel distribution.

Usage

```
f.gumb (x, xi, alfa)
F.gumb (x, xi, alfa)
invF.gumb (F, xi, alfa)
Lmom.gumb (xi, alfa)
par.gumb (lambda1, lambda2)
rand.gumb (numerosita, xi, alfa)
```

Arguments

x	vector of quantiles
xi	vector of gumb location parameters
alfa	vector of gumb scale parameters
F	vector of probabilities
lambda1	vector of sample means
lambda2	vector of L-variances
numerosita	numeric value indicating the length of the vector to be generated

Details

See http://en.wikipedia.org/wiki/Fisher-Tippett_distribution for an introduction to the Gumbel distribution.

Definition

Parameters (2): ξ (location), α (scale).

Range of x: $-\infty < x < \infty$.

Probability density function:

$$f(x) = \alpha^{-1} \exp[-(x - \xi)/\alpha] \exp\{-\exp[-(x - \xi)/\alpha]\}$$

Cumulative distribution function:

$$F(x) = \exp[-\exp(-(x - \xi)/\alpha)]$$

Quantile function: $x(F) = \xi - \alpha \log(-\log F)$.

L-moments

$$\lambda_1 = \xi + \alpha\gamma$$

$$\lambda_2 = \alpha \log 2$$

$$\tau_3 = 0.1699 = \log(9/8)/\log 2$$

$$\tau_4 = 0.1504 = (16 \log 2 - 10 \log 3)/\log 2$$

Here γ is Euler's constant, 0.5772...

Parameters

$$\alpha = \lambda_2/\log 2$$

$$\xi = \lambda_1 - \gamma\alpha$$

Value

f.gumb gives the density f, F.gumb gives the distribution function F, invF.gumb gives the quantile function x, Lmom.gumb gives the L-moments (λ_1, λ_2, τ_3, τ_4)), par.gumb gives the parameters (xi, alfa), and rand.gumb generates random deviates.

Note

Lmom.gumb and par.gumb accept input as vectors of equal length. In f.gumb, F.gumb, invF.gumb and rand.gumb parameters (xi, alfa) must be atomic.

Author(s)

Alberto Viglione, e-mail: <alviglio@tiscali.it>.

References

Hosking, J.R.M. and Wallis, J.R. (1997) Regional Frequency Analysis: an approach based on L-moments, Cambridge University Press, Cambridge, UK.

See Also

rnorm, runif, GEV, Lmoments.

HOMTESTS   *Homogeneity tests*

Description

Homogeneity tests for Regional Frequency Analysis.

Usage

```
ADbootstrap.test (x, cod, Nsim=500, index=2)
HW.tests (x, cod, Nsim=500)
DK.test (x, cod)
```

Arguments

x	vector representing data from many samples defined with cod
cod	array that defines the data subdivision among sites
Nsim	number of regions simulated with the bootstrap of the original region
index	if index=1 samples are divided by their average value; if index=2 (default) samples are divided by their median value

Details

The Hosking and Wallis heterogeneity measures

The idea underlying Hosking and Wallis (1993) heterogeneity statistics is to measure the sample variability of the L-moment ratios and compare it to the variation that would be expected in a homogeneous region. The latter is estimated through repeated simulations of homogeneous regions with samples drawn from a four parameter kappa distribution (see e.g., Hosking and Wallis, 1997, pp. 202-204). More in detail, the steps are the following: with regards to the k samples belonging to the region under analysis, find the sample L-moment ratios (see, Hosking and Wallis, 1997) pertaining to the i-th site: these are the L-coefficient of variation (L-CV),

$$t^{(i)} = \frac{\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{2(j-1)}{(n_i-1)} - 1 \right) Y_{i,j}}{\frac{1}{n_i} \sum_{j=1}^{n_i} Y_{i,j}}$$

the coefficient of L-skewness,

$$t_3^{(i)} = \frac{\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{6(j-1)(j-2)}{(n_i-1)(n_i-2)} - \frac{6(j-1)}{(n_i-1)} + 1 \right) Y_{i,j}}{\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{2(j-1)}{(n_i-1)} - 1 \right) Y_{i,j}}$$

and the coefficient of L-kurtosis

$$t_4^{(i)} = \frac{\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{20(j-1)(j-2)(j-3)}{(n_i-1)(n_i-2)(n_i-3)} - \frac{30(j-1)(j-2)}{(n_i-1)(n_i-2)} + \frac{12(j-1)}{(n_i-1)} - 1 \right) Y_{i,j}}{\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{2(j-1)}{(n_i-1)} - 1 \right) Y_{i,j}}$$

Note that the L-moment ratios are not affected by the normalization by the index value, i.e. it is the same to use $X_{i,j}$ or $Y_{i,j}$ in Equations.

Define the regional averaged L-CV, L-skewness and L-kurtosis coefficients,

$$t^R = \frac{\sum_{i=1}^{k} n_i t^{(i)}}{\sum_{i=1}^{k} n_i}$$

$$t_3^R = \frac{\sum_{i=1}^{k} n_i t_3^{(i)}}{\sum_{i=1}^{k} n_i}$$

$$t_4^R = \frac{\sum_{i=1}^{k} n_i t_4^{(i)}}{\sum_{i=1}^{k} n_i}$$

and compute the statistic

$$V = \left\{ \sum_{i=1}^{k} n_i (t^{(i)} - t^R)^2 / \sum_{i=1}^{k} n_i \right\}^{1/2}$$

Fit the parameters of a four-parameters kappa distribution to the regional averaged L-moment ratios t^R, t_3^R and t_4^R, and then generate a large number N_{sim} of realizations of sets of k samples. The i-th site sample in each set has a kappa distribution as its parent and record length equal to n_i. For each simulated homogeneous set, calculate the statistic V, obtaining N_{sim} values. On this vector of V values determine the mean μ_V and standard deviation σ_V that relate to the hypothesis of homogeneity (actually, under the composite hypothesis of homogeneity and kappa parent distribution).

An heterogeneity measure, which is called here HW_1, is finally found as

$$\theta_{HW_1} = \frac{V - \mu_V}{\sigma_V}$$

θ_{HW_1} can be approximated by a normal distributed with zero mean and unit variance: following Hosking and Wallis (1997), the region under analysis can therefore be regarded as 'acceptably homogeneous' if $\theta_{HW_1} < 1$, 'possibly heterogeneous' if $1 \leq \theta_{HW_1} < 2$, and 'definitely heterogeneous' if $\theta_{HW_1} \geq 2$. Hosking and Wallis (1997) suggest that these limits should be treated as useful guidelines. Even if the θ_{HW_1} statistic is constructed like a significance test, significance levels obtained from such a test would in fact be accurate only under special assumptions: to have independent data both serially and between sites, and the true regional distribution being kappa.

Hosking and Wallis (1993) also give an alternative heterogeneity measure (that we call HW_2), in which V is replaced by:

$$V_2 = \sum_{i=1}^{k} n_i \left\{ (t^{(i)} - t^R)^2 + (t_3^{(i)} - t_3^R)^2 \right\}^{1/2} / \sum_{i=1}^{k} n_i$$

The test statistic in this case becomes

$$\theta_{HW_2} = \frac{V_2 - \mu_{V_2}}{\sigma_{V_2}}$$

with similar acceptability limits as the HW_1 statistic. Hosking and Wallis (1997) judge θ_{HW_2} to be inferior to θ_{HW_1} and say that it rarely yields values larger than 2 even for grossly heterogeneous regions.

The bootstrap Anderson-Darling test

A test that does not make any assumption on the parent distribution is the Anderson-Darling (AD) rank test (Scholz and Stephens, 1987). The AD test is the generalization of the classical Anderson-Darling goodness of fit test (e.g., D'Agostino and Stephens, 1986), and it is used to test the hypothesis that k independent samples belong to the same population without specifying their common distribution function.

The test is based on the comparison between local and regional empirical distribution functions. The empirical distribution function, or sample distribution function, is defined by $F(x) = \frac{j}{\eta}, x_{(j)} \leq x < x_{(j+1)}$, where η is the size of the sample and $x_{(j)}$ are the order statistics, i.e. the observations arranged in ascending order. Denote the empirical distribution function of the i-th sample (local) by $\hat{F}_i(x)$, and that of the pooled sample of all $N = n_1 + ... + n_k$ observations (regional) by $H_N(x)$. The k-sample Anderson-Darling test statistic is then defined as

$$\theta_{AD} = \sum_{i=1}^{k} n_i \int_{\text{all } x} \frac{[\hat{F}_i(x) - H_N(x)]^2}{H_N(x)[1 - H_N(x)]} dH_N(x)$$

If the pooled ordered sample is $Z_1 < ... < Z_N$, the computational formula to evaluate θ_{AD} is:

$$\theta_{AD} = \frac{1}{N} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{N-1} \frac{(NM_{ij} - jn_i)^2}{j(N - j)}$$

where M_{ij} is the number of observations in the i-th sample that are not greater than Z_j. The homogeneity test can be carried out by comparing the obtained θ_{AD} value to the tabulated percentage points reported by Scholz and Stephens (1987) for different significance levels.

The statistic θ_{AD} depends on the sample values only through their ranks. This guarantees that the test statistic remains unchanged when the samples undergo monotonic transformations, an important stability property not possessed by HW heterogeneity measures. However, problems arise in applying this test in a common index value procedure. In fact, the index value procedure corresponds to dividing each site sample by a different value, thus modifying the ranks in the pooled sample. In particular, this has the effect of making the local empirical distribution functions much more similar to the other, providing an impression of homogeneity even when the samples are highly heterogeneous. The effect is analogous to that encountered when applying goodness-of-fit tests to distributions whose parameters are estimated from the same sample used for the test (e.g., D'Agostino and Stephens, 1986; Laio, 2004). In both cases, the percentage points for the test should be opportunely redetermined. This can be done with a nonparametric bootstrap approach presenting the following steps: build up the pooled sample $\mathcal{S}$ of the observed non-dimensional data. Sample with replacement from $\mathcal{S}$ and generate k artificial local samples, of size $n_1, \ldots, n_k$. Divide each sample for its index value, and calculate $\theta_{AD}^{(1)}$. Repeat the procedure for N_{sim} times and obtain a sample of $\theta_{AD}^{(j)}$, $j = 1, \ldots, N_{sim}$ values, whose empirical distribution function can be used as an approximation of $G_{H_0}(\theta_{AD})$, the distribution of θ_{AD} under the null hypothesis of homogeneity. The acceptance limits for the test, corresponding to any significance level α, are then easily determined as the quantiles of $G_{H_0}(\theta_{AD})$ corresponding to a probability $(1 - \alpha)$.

We will call the test obtained with the above procedure the bootstrap Anderson-Darling test, hereafter referred to as AD.

Durbin and Knott test

The last considered homogeneity test derives from a goodness-of-fit statistic originally proposed by Durbin and Knott (1971). The test is formulated to measure discrepancies in the dispersion of the samples, without accounting for the possible presence of discrepancies in the mean or skewness of the data. Under this aspect, the test is similar to the HW_1 test, while it is analogous to the AD test for the fact that it is a rank test. The original goodness-of-fit test is very simple: suppose to have a sample X_i, $i = 1, ..., n$, with hypothetical distribution $F(x)$; under the null hypothesis the random variable $F(X_i)$ has a uniform distribution in the $(0, 1)$ interval, and the statistic $D = \sum_{i=1}^{n} \cos[2\pi F(X_i)]$ is approximately normally distributed with mean 0 and variance 1 (Durbin and Knott, 1971). D serves the purpose of detecting discrepancy in data dispersion: if the variance of X_i is greater than that of the hypothetical distribution $F(x)$, D is significantly greater than 0, while D is significantly below 0 in the reverse case. Differences between the mean (or the median) of X_i and $F(x)$ are instead not detected by D, which guarantees that the normalization by the index value does not affect the test.

The extension to homogeneity testing of the Durbin and Knott (DK) statistic is straightforward: we substitute the empirical distribution function obtained with the pooled observed data, $H_N(x)$, for $F(x)$ in D, obtaining at each site a statistic

$$D_i = \sum_{j=1}^{n_i} \cos[2\pi H_N(X_j)]$$

which is normal under the hypothesis of homogeneity. The statistic $\theta_{DK} = \sum_{i=1}^{k} D_i^2$ has then a chi-squared distribution with $k - 1$ degrees of freedom, which allows one to determine the acceptability limits for the test, corresponding to any significance level α.

Comparison among tests

The comparison (Viglione et al, 2007) shows that the Hosking and Wallis heterogeneity measure HW_1 (only based on L-CV) is preferable when skewness is low, while the bootstrap Anderson-

Darling test should be used for more skewed regions. As for HW_2, the Hosking and Wallis heterogeneity measure based on L-CV and L-CA, it is shown once more how much it lacks power.

Our suggestion is to guide the choice of the test according to a compromise between power and Type I error of the HW_1 and AD tests. The L-moment space is divided into two regions: if the t_3^R coefficient for the region under analysis is lower than 0.23, we propose to use the Hosking and Wallis heterogeneity measure HW_1; if $t_3^R > 0.23$, the bootstrap Anderson-Darling test is preferable.

Value

`ADbootstrap.test` and `DK.test` test gives its test statistic and its distribution value P. If P is, for example, 0.92, samples shouldn't be considered heterogeneous with significance level minor of 8

`HW.tests` gives the two Hosking and Wallis heterogeneity measures HW_1 and HW_2; following Hosking and Wallis (1997), the region under analysis can therefore be regarded as 'acceptably homogeneous' if $HW < 1$, 'possibly heterogeneous' if $1 \leq HW < 2$, and 'definitely heterogeneous' if $HW \geq 2$.

Author(s)

Alberto Viglione, e-mail: `<alviglio@tiscali.it>`.

References

D'Agostino R., Stephens M. (1986) Goodness-of-Fit Techniques, chapter Tests based on EDF statistics. Marcel Dekker, New York.

Durbin J., Knott M. (1971) Components of Cramer-von Mises statistics. London School of Economics and Political Science, pp. 290-307.

Hosking J., Wallis J. (1993) Some statistics useful in regional frequency analysis. Water Resources Research, 29 (2), pp. 271-281.

Hosking, J.R.M. and Wallis, J.R. (1997) Regional Frequency Analysis: an approach based on L-moments, Cambridge University Press, Cambridge, UK.

Laio, F., Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, Water Resour. Res., 40, W09308, doi:10.1029/2004WR003204.

Scholz F., Stephens M. (1987) K-sample Anderson-Darling tests. Journal of American Statistical Association, 82 (399), pp. 918-924.

Viglione A., Laio F., Claps P. (2007) "A comparison of homogeneity tests for regional frequency analysis", Water Resources Research, 43, W03428, doi:10.1029/2006WR005095.

Viglione A. (2007) Metodi statistici non-supervisati per la stima di grandezze idrologiche in siti non strumentati, PhD thesis, Politecnico di Torino.

See Also

KAPPA, Lmoments.

Examples

```
data(annualflows)
annualflows[1:10,]
summary(annualflows)
x <- annualflows["dato"][,]
cod <- annualflows["cod"][,]
split(x,cod)

#ADbootstrap.test(x,cod,Nsim=100)   # it takes some time
#HW.tests(x,cod)                     # it takes some time
DK.test(x,cod)

fac <- factor(annualflows["cod"][,],levels=c(34:38))
x2 <- annualflows[!is.na(fac),"dato"]
cod2 <- annualflows[!is.na(fac),"cod"]
split(x2,cod2)
sapply(split(x2,cod2),Lmoments)
regionalLmoments(x2,cod2)

ADbootstrap.test(x2,cod2)
ADbootstrap.test(x2,cod2,index=1)
HW.tests(x2,cod2)
DK.test(x2,cod2)
```

KAPPA	*Four parameter kappa distribution and L-moments*

Description

KAPPA provides the link between L-moments of a sample and the four parameter kappa distribution.

Usage

```
f.kappa (x, xi, alfa, k, h)
F.kappa (x, xi, alfa, k, h)
invF.kappa (F, xi, alfa, k, h)
Lmom.kappa (xi, alfa, k, h)
par.kappa (lambda1, lambda2, tau3, tau4)
rand.kappa (numerosita, xi, alfa, k, h)
```

Arguments

x	vector of quantiles
xi	vector of kappa location parameters
alfa	vector of kappa scale parameters
k	vector of kappa third parameters
h	vector of kappa fourth parameters

F	vector of probabilities
lambda1	vector of sample means
lambda2	vector of L-variances
tau3	vector of L-CA (or L-skewness)
tau4	vector of L-kurtosis
numerosita	numeric value indicating the length of the vector to be generated

Details

Definition

Parameters (4): ξ (location), α (scale), k, h.

Range of x: upper bound is $\xi + \alpha/k$ if $k > 0$, ∞ if $k \leq 0$; lower bound is $\xi + \alpha(1 - h^{-k})/k$ if $h > 0$, $\xi + \alpha/k$ if $h \leq 0$ and $k < 0$ and $-\infty$ if $h \leq 0$ and $k \geq 0$

Probability density function:

$$f(x) = \alpha^{-1}[1 - k(x - \xi)/\alpha]^{1/k-1}[F(x)]^{1-h}$$

Cumulative distribution function:

$$F(x) = \{1 - h[1 - k(x - \xi)/\alpha]^{1/k}\}^{1/h}$$

Quantile function:

$$x(F) = \xi + \frac{\alpha}{k}\left[1 - \left(\frac{1 - F^h}{h}\right)^k\right]$$

$h = -1$ is the generalized logistic distribution; $h = 0$ is the generalized eztreme value distribution; $h = 1$ is the generalized Pareto distribution.

L-moments

L-moments are defined for $h \geq 0$ and $k > -1$, or if $h < 0$ and $-1 < k < -1/h$.

$$\lambda_1 = \xi + \alpha(1 - g_1)/k$$

$$\lambda_2 = \alpha(g_1 - g_2)/k$$

$$\tau_3 = (-g_1 + 3g_2 - 2g_3)/(g_1 - g_2)$$

$$\tau_4 = (-g_1 + 6g_2 - 10g_3 + 5g_4)/(g_1 - g_2)$$

where $g_r = \frac{r\Gamma(1+k)\Gamma(r/h)}{h^{1+k}\Gamma(1+k+r/h)}$ if $h > 0$; $g_r = \frac{r\Gamma(1+k)\Gamma(-k-r/h)}{(-h)^{1+k}\Gamma(1-r/h)}$ if $h < 0$;

Here Γ denote the gamma function

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

Parameters

There are no simple expressions for the parameters in terms of the L-moments. However they can be obtained with a numerical algorithm considering the formulations of τ_3 and τ_4 in terms of k and h. Here we use the function `optim` to minimize $(t_3 - \tau_3)^2 + (t_4 - \tau_4)^2$ where t_3 and t_4 are the sample L-moment ratios.

Value

f.kappa gives the density f, F.kappa gives the distribution function F, invFkappa gives the quantile function x, Lmom.kappa gives the L-moments (λ_1, λ_2, τ_3, τ_4), par.kappa gives the parameters (xi, alfa, k, h), and rand.kappa generates random deviates.

Note

Lmom.kappa and par.kappa accept input as vectors of equal length. In f.kappa, F.kappa, invF.kappa and rand.kappa parameters (xi, alfa, k, h) must be atomic.

Author(s)

Alberto Viglione, e-mail: <alviglio@tiscali.it>.

References

Hosking, J.R.M. and Wallis, J.R. (1997) Regional Frequency Analysis: an approach based on L-moments, Cambridge University Press, Cambridge, UK.

See Also

HOMTESTS, rnorm, runif.

Examples

```
data(annualflows)
annualflows
summary(annualflows)
x <- annualflows["dato"][,]
fac <- factor(annualflows["cod"][,])
split(x,fac)

camp <- split(x,fac)$"45"
ll <- Lmoments(camp)
parameters <- par.kappa(ll[1],ll[2],ll[4],ll[5])
f.kappa(1800,parameters$xi,parameters$alfa,parameters$k,parameters$h)
F.kappa(1800,parameters$xi,parameters$alfa,parameters$k,parameters$h)
invF.kappa(0.771088,parameters$xi,parameters$alfa,parameters$k,parameters$h)
Lmom.kappa(parameters$xi,parameters$alfa,parameters$k,parameters$h)
rand.kappa(100,parameters$xi,parameters$alfa,parameters$k,parameters$h)

Rll <- regionalLmoments(x,fac); Rll
parameters <- par.kappa(Rll[1],Rll[2],Rll[4],Rll[5])
Lmom.kappa(parameters$xi,parameters$alfa,parameters$k,parameters$h)
```

Lmoments *Hosking and Wallis sample L-moments*

Description

Lmoments provides the estimate of L-moments of a sample or regional L-moments of a region.

Usage

```
Lmoments (x)
regionalLmoments (x,cod)
LCV (x)
LCA (x)
Lkur (x)
```

Arguments

x	vector representing a data-sample (or data from many samples defined with cod in the case of regionalLmoments)
cod	array that defines the data subdivision among sites

Details

The estimation of L-moments is based on a sample of size n, arranged in ascending order. Let $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$ be the ordered sample. An unbiased estimator of the probability weighted moments β_r is:

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_{j:n}$$

The sample L-moments are defined by:
$$l_1 = b_0$$

$$l_2 = 2b_1 - b_0$$

$$l_3 = 6b_2 - 6b_1 + b_0$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$$

and in general

$$l_{r+1} = \sum_{k=0}^{r} \frac{(-1)^{r-k}(r+k)!}{(k!)^2(r-k)!} b_k$$

where $r = 0, 1, \ldots, n-1$.

The sample L-moment ratios are defined by

$$t_r = l_r/l_2$$

and the sample L-CV by

$$t = l_2/l_1$$

Sample regional L-CV, L-skewness and L-kurtosis coefficients are defined as

$$t^R = \frac{\sum_{i=1}^{k} n_i t^{(i)}}{\sum_{i=1}^{k} n_i}$$

$$t_3^R = \frac{\sum_{i=1}^{k} n_i t_3^{(i)}}{\sum_{i=1}^{k} n_i}$$

$$t_4^R = \frac{\sum_{i=1}^{k} n_i t_4^{(i)}}{\sum_{i=1}^{k} n_i}$$

Value

`Lmoments` gives the L-moments (l_1, l_2, t, t_3, t_4), `regionalLmoments` gives the regional weighted L-moments (l_1^R, l_2^R, t^R, t_3^R, t_4^R), `LCV` gives the coefficient of L-variation, `LCA` gives the L-skewness and `Lkur` gives the L-kurtosis of `x`.

Author(s)

Alberto Viglione, e-mail: <alviglio@tiscali.it>.

References

Hosking, J.R.M. and Wallis, J.R. (1997) Regional Frequency Analysis: an approach based on L-moments, Cambridge University Press, Cambridge, UK.

See Also

mean, var, sd, HOMTESTS.

Examples

```
x <- rnorm(30,10,2)
Lmoments(x)

data(annualflows)
annualflows
summary(annualflows)
x <- annualflows["dato"][,]
cod <- annualflows["cod"][,]
split(x,cod)
camp <- split(x,cod)$"45"
Lmoments(camp)
sapply(split(x,cod),Lmoments)

regionalLmoments(x,cod)
```

Index