Package 'lmomRFA'

May 29, 2022

Version 3.4

Date 2022-05-29

Title Regional Frequency Analysis using L-Moments

Author J. R. M. Hosking

Maintainer J. R. M. Hosking <jrmhosking@gmail.com>

Description Functions for regional frequency analysis using the methods
of J. R. M. Hosking and J. R. Wallis (1997), ``Regional frequency analysis:
an approach based on L-moments".

Depends R (>= 3.0.0), lmom (>= 2.0)

Imports stats, graphics

License Common Public License Version 1.0

LazyData yes

Archs x64

R topics documented:

---

lmomRFA-package   *The lmomRFA package*

Description

R functions for regional frequency analysis using L-moments.

Details

This package implements methods described in the book "Regional frequency analysis: an approach based on L-moments" by J. R. M. Hosking and J. R. Wallis. It is a supplement to the **lmom** package, which implements L-moment methods for more general statistical applications.

The following functions are contained in this package.

cluagg performs agglomerative hierarchical clustering.

cluinf provides information about cluster membership in a hierarchical clustering.

clukm performs cluster analysis via the K-means algorithm.

as.regdata creates an object of class "regdata", which contains a "regional data set" consisting of summary statistics for different data samples, one of the main building blocks of regional frequency analysis.

regsamlmu computes the sample L-moments of multiple data sets.

regavlmom and reglmr both compute, with slightly different interfaces, a regional weighted average of sample L-moments from multiple sites. Function regavlmom is recommended for general use; reglmr is deprecated.

regtst computes discordancy, heterogeneity and goodness-of-fit measures for regional frequency analysis. These statistics are as described in Hosking and Wallis (1997, chaps. 3-5).

regfit fits a frequency distribution to a regional data set, giving a "regional frequency distribution".

regqfunc and siteqfunc return the regional growth curve and the quantile functions for individual sites, respectively, from a regional frequency distribution fitted by regfit.

regquant and sitequant directly compute quantiles of the regional growth curve and of distributions for individual sites, respectively, from a regional frequency distribution fitted by regfit.

regsimh runs Monte Carlo simulations to estimate the distribution of heterogeneity and goodness-of-fit measures for an artificial region.

regsimq runs Monte Carlo simulations to estimate the variability of quantile estimates from a regional frequency distribution.

regquantbounds and sitequantbounds compute error bounds for the regional growth curve and for quantiles at individual sites, respectively, from a regional frequency distribution fitted by regfit.

Functions cluagg, cluinf, clukm, reglmr, and regtst are analogous to Fortran routines from the LMOMENTS package, version 3.04, available from StatLib at http://lib.stat.cmu.edu/general/lmoments. In addition, functions regsimh and regsimq provide similar functionality to PROGRAM XSIM in the LMOMENTS Fortran package.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

Appalach	*Data for streamflow gaging stations in Appalachia*

Description

Site characteristics and sample L-moments of annual maximum streamflow for 104 gaging stations in Appalachia.

Usage

```
Appalach
```

Format

A data frame with 104 observations on the following 11 variables:

siteid Character vector: each site's Hydrologic Unit Code, a unique identifier.

lat Numeric vector: gage latitude, in degrees.

long Numeric vector: gage longitude, in degrees west of the Greenwich Meridian.

area Numeric vector: drainage basin area, in square miles.

elev Numeric vector: gage elevation, in feet.

n Numeric vector: record length.

mean Numeric vector: sample mean.

t Numeric vector: sample L-CV.

t_3 Numeric vector: sample L-skewness.

t_4 Numeric vector: sample L-kurtosis.

t_5 Numeric vector: sample L-moment ratio t_5.

Details

The data in columns lat, long, area and elev, and the streamflow data used to compute the sample L-moments, were obtained from "Hydrodata" CD-ROMs (Hydrosphere, 1993), which reproduce data from the U.S. Geological Survey's WATSTORE data files.

Source

The file appalach.dat in the LMOMENTS Fortran package (Hosking, 1996).

References

Hydrosphere (1993). *Hydrodata CD-ROMs, vol. 4.0: USGS peak values*. Hydrosphere Data Products, Boulder, Colo.

Hosking, J. R. M. (1996). Fortran routines for use with the method of L-moments, Version 3. Research Report RC20525, IBM Research Division, Yorktown Heights, N.Y.

Examples

    Appalach

Cascades	*L-moments of annual precipitation totals*

Description

L-moments of annual precipitation totals for the "North Cascades" region of Plantico et al. (1990).

Usage

    Cascades

Format

An object of class `regdata`. It is a data frame with 19 observations on the following 7 variables:

name Character vector: site identifier.
n Numeric vector: record length.
mean Numeric vector: sample mean.
t Numeric vector: sample *L*-CV.
t_3 Numeric vector: sample *L*-skewness.
t_4 Numeric vector: sample *L*-kurtosis.
t_5 Numeric vector: sample *L*-moment ratio t_5.

Details

The data are summary statistics of annual precipitation totals at 19 sites in the northwest U.S., the "North Cascades" region of Plantico et al. (1990). The precipitation data were obtained from the Historical Climatology Network (Karl et al., 1990).

Source

The file `cascades.dat` in the LMOMENTS Fortran package (Hosking, 1996).

References

Hosking, J. R. M. (1996). Fortran routines for use with the method of *L*-moments, Version 3. Research Report RC20525, IBM Research Division, Yorktown Heights, N.Y.

Karl, T. R., Williams, C. N., Quinlan, F. T., and Boden, T. A. (1990). United States historical climatology network (HCN) serial temperature and precipitation data. *ORNL/CDIAC-30 NDP-019/R1*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn.

Plantico, M. S., Karl, T. R., Kukla, G., and Gavin, J. (1990). Is recent climate change across the United States related to rising levels of anthropogenic greenhouse gases? *Journal of Geophysical Research*, **95**, 16617–16637.

Examples

    Cascades   # L-moment ratios, etc., for the Cascades data
    lmrd(Cascades) # L-moment ratio diagram for the Cascades data

---

cluagg	*Hierarchical clustering*

Description

Performs cluster analysis by one of several agglomerative hierarchical methods.

Usage

```
cluagg(x, method="ward")
```

Arguments

x
: A numeric matrix (or a data frame with all numeric columns, which will be coerced to a matrix). Contains the data: each row should contain the attributes for a single point.

method
: Clustering method. Any method valid for hclust may be used.

Details

In agglomerative hierarchical clustering, there are initially n clusters, each containing one data point, labeled 1 through n in the same order as the data points. At each stage of clustering, two clusters are merged. Their labels are saved in the merge array. The smaller of the two labels is used as the label of the merged cluster. After the ith stage of clustering there are $n - i$ clusters. To find which data points belong to which clusters, use function cluinf.

Value

A list with elements as follows.

merge
: Matrix of dimension (nrow(x)-1,2). The ith row contains the labels of the clusters merged at the ith merge.

wgss
: Vector of length nrow(x)-1. The ith element is the total within-cluster dispersion after the ith merge.

Note

Clustering is performed internally by function hclust in the R **stats** package.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

cluinf to get details of the clusters at a particular stage of the merging.

Examples

```
## Clustering of gaging stations in Appalachia, as in Hosking
## and Wallis (1997, sec. 9.2.3)
data(Appalach)
# Form attributes for clustering (Hosking and Wallis's Table 9.4)
att <- cbind(a1 = log(Appalach$area),
             a2 = sqrt(Appalach$elev),
             a3 = Appalach$lat,
             a4 = Appalach$long)
att <- apply(att, 2, function(x) x/sd(x))
att[,1] <- att[,1] * 3
# Clustering by Ward's method
(cl<-cluagg(att))
# Details of the clustering with 7 clusters
cluinf(cl,7)
```

---

cluinf   *Provide information about a hierarchical clustering*

Description

Agglomerative hierarchical clustering procedures typically produce a list of the clusters merged at
each stage of the clustering. `cluinf` uses this list to construct arrays that explicitly show which
cluster a given data point belongs to, and which data points belong to a given cluster.

Usage

```
cluinf(merge, nclust)
```

Arguments

merge	Matrix with 2 columns. The ith row contains the labels of the clusters merged at the ith merge.
	Can also be the object returned by a call to cluagg.
nclust	Number of clusters.

Value

Information about the clustering that has `nclust` clusters. It is a list with the following elements:

assign	Vector giving the assignment of items to clusters.
list	List with `nclust` elements. Each element contains the labels of the items in one cluster.
num	Vector of length `nclust`, containing the number of items in each cluster.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on
L-moments*. Cambridge University Press.

See Also

[cluagg](#)

Examples

```
## Clustering of gaging stations in Appalachia, as in Hosking
## and Wallis (1997, sec. 9.2.3)
data(Appalach)
# Form attributes for clustering (Hosking and Wallis's Table 9.4)
att <- cbind(a1 = log(Appalach$area),
             a2 = sqrt(Appalach$elev),
             a3 = Appalach$lat,
             a4 = Appalach$long)
att <- apply(att, 2, function(x) x/sd(x))
att[,1] <- att[,1] * 3
# Clustering by Ward's method
(cl<-cluagg(att))
# Details of the clustering with 7 clusters
cluinf(cl, 7)
```

---

clukm	*Cluster analysis via K-means algorithm*

Description

Performs cluster analysis using the K-means algorithm.

Usage

```
clukm(x, assign, maxit = 10, algorithm = "Hartigan-Wong")
```

Arguments

x	A numeric matrix (or a data frame with all numeric columns, which will be coerced to a matrix). Contains the data: each row should contain the attributes for a single point.
assign	A vector whose distinct values indicate the initial clustering of the points.
maxit	Maximum number of iterations.
algorithm	Clustering algorithm. Permitted values are the same as for kmeans.

Value

An object of class kmeans. For details see the help for [kmeans](#).

Note

clukm is a wrapper for the R function kmeans. The only difference is that in clukm the user supplies an initial assignment of sites to clusters (from which cluster centers are computed), whereas in kmeans the user supplies the initial cluster centers explicitly.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

kmeans

Examples

```
## Clustering of gaging stations in Appalachia, as in Hosking
## and Wallis (1997, sec. 9.2.3)
data(Appalach)
# Form attributes for clustering (Hosking and Wallis's Table 9.4)
att <- cbind(a1 = log(Appalach$area),
             a2 = sqrt(Appalach$elev),
             a3 = Appalach$lat,
             a4 = Appalach$long)
att <- apply(att, 2, function(x) x/sd(x))
att[,1] <- att[,1] * 3
# Clustering by Ward's method
(cl <- cluagg(att))
# Details of the clustering with 7 clusters
(inf <- cluinf(cl, 7))
# Refine the 7 clusters by K-means
clkm <- clukm(att, inf$assign)
# Compare the original and K-means clusters
table(Kmeans=clkm$cluster, Ward=inf$assign)
# Some details about the K-means clusters: range of area, number
# of sites, weighted average L-CV and L-skewness
bb <- by(Appalach, clkm$cluster, function(x)
  c( min.area = min(x$area),
     max.area = max(x$area),
     n = nrow(x),
     ave.t = round(weighted.mean(x$t, x$n), 3),
     ave.t_3 = round(weighted.mean(x$t_3, x$n), 3)))
# Order the clusters in increasing order of minimum area
ord <- order(sapply(bb, "[", "min.area"))
# Make the result into a data frame.  Compare with Hosking
# and Wallis (1997), Table 9.5.
do.call(rbind, bb[ord])
```

---

evplot.rfd   *Extreme-value plot of a regional frequency distribution*

Description

Plots a regional frequency distribution, optionally with error bounds for either the regional growth curve or the quantile function for an individual site.

The graph is an "extreme-value plot", i.e. the horizontal axis is the quantile of an extreme-value type I (Gumbel) distribution, and the quantile function of that distribution would plot as a straight line.

Usage

```
## S3 method for class 'rfd'
evplot(y, ybounds, npoints=101, add=FALSE, plim,
  xlim=c(-2,5), ylim,
  xlab=expression("Reduced variate,  " * -log(-log(italic(F)))),
  ylab="Quantile", rp.axis=TRUE, type="l", lty=c(1,2), col=c(1,1),
  ...)
```

Arguments

y	Object of class [rfd](), containing the specification of a regional frequency distribution.
ybounds	Optional. Object of class rfdbounds (typically created by [regquantbounds]() or [sitequantbounds]()), containing error bounds for quantile estimates for the regional frequency distribution specified by y.
npoints	Number of points to use in drawing the quantile function. The points are equally spaced along the x axis.
add	Logical: if TRUE, add to existing plot.
plim	X axis limits, specified as probabilities.
xlim	X axis limits, specified as values of the Gumbel reduced variate $-\log(-\log(F))$, where F is the nonexceedance probability. Not used if plim is specified.
ylim	Y axis limits.
xlab	X axis label.
ylab	Y axis label.
rp.axis	Logical: whether to draw the "Return period" axis, a secondary horizontal axis.
type	Vector of plot types. The first element is for the quantile function; subsequent elements are for the error bounds, and will be used cyclically until all lines are drawn. Interpreted in the same way as the type plotting parameter, i.e. "l" for lines, "b" for points connected by lines, etc.
lty	Vector of line types. The first element is for the quantile function; subsequent elements are for the error bounds, and will be used cyclically until all lines are drawn.
col	Vector of colors. The first element is for the quantile function; subsequent elements are for the error bounds, and will be used cyclically until all lines are drawn.
...	Additional parameters are passed to the plotting routine.

Details

If ybounds is missing, a graph is drawn of the quantile function (regional growth curve) of the distribution specified by y.

If ybounds is present, it may contain error bounds for either a regional growth curve or the quantile function at a single site. This regional growth curve or site quantile function is plotted using arguments type[1], lty[1] and col[1]. Then, in each case, error bounds are added to the plot. The ybounds object typically contains, for several probabilities specified by ybounds$bounds, error bounds corresponding to that probability for several quantiles. For the jth bound probability, the bounds for the various quantiles will be joined by straight lines (so to obtain a smooth curve there should be a lot of quantiles!), using graphics parameters type[j+1], lty[j+1] and col[j+1].

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

See Also

`regfit`, which creates objects of class `"rfd"`; `regquantbounds` and `sitequantbounds`, which create objects of class `"rfdbounds"`;

evdistp, evdistq, and evpoints, all in package **lmom**, for adding further curves and points to the plot.

Examples

```
Cascades                       # An object of class "regdata"
rfit <- regfit(Cascades, "gno") # Fit a generalized normal distribution

evplot(rfit)                   # Plot the regional growth curve

# Compute error bounds for quantile estimates.  We will
# (optimistically) generate bounds for a homogeneous region
# with the same frequency distribution as the one fitted to
# the Cascades data.
fval <- seq(.01, .99, by=.01)   # A lot of quantiles
simq <- regsimq(rfit$qfunc, nrec=Cascades$n, nrep=100, f=fval,
  fit=rfit$dist)

# Regional growth curve, and bounds
rbounds <- regquantbounds(simq, rfit)
evplot(rfit, rbounds)

# Quantile function for site 3, and bounds
sbounds <- sitequantbounds(simq, rfit, site=3)
evplot(rfit, sbounds)
```

Maxwind   *Maximum wind speeds*

Description

Annual maximum wind speeds at 12 sites in the southeast U.S.

Usage

```
Maxwind
```

Format

A list of 12 numeric vectors.

Details

The name of a list element is the site location, including a reference number used by Simiu et al. (1979). Each list element is a numeric vector containing the annual maximum wind speeds for that site. The period of observation varies from site to site: for details see Simiu et al. (1979).

Source

Simiu, E., Changery, M. J., and Filliben, J. J. (1979). Extreme wind speeds at 129 stations in the contiguous United States. *Building Science Series 118*, National Bureau of Standards, Washington, D.C.

Examples

```
str(Maxwind)
regsamlmu(Maxwind)  # sample L-moments
```

---

regavlmom	*Regional weighted average of L-moments*

Description

Computes a regional weighted average of *L*-moments.

Usage

```
regavlmom(regdata, weight)
```

Arguments

regdata   Object of class regdata, containing summary statistics of the data for the sites in a region.

weight   Vector containing the weights to be used for each site. If omitted, weights will be the sample size at each site, taken from the second column of regdata. If a single value, equal weights will be used.

Value

Vector containing the regional average *L*-moments.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

Examples

```
reglm <- regsamlmu(Maxwind)
regavlmom(reglm)              # Weight proportional to record length
regavlmom(reglm, weight=1)  # Equal weights
```

---

regdata   *The regdata class*

Description

An object of class `"regdata"` stores summary statistics of the data for the sites in a region. It is a data frame with each row containing data for one site. The columns should contain the site name, record length and L-moments and L-moment ratios, in the order ℓ_1 (mean), t (L-CV), t_3 (L-skewness), t_4 (L-kurtosis), t_5, t_6, etc.

There should be at least four columns, but most functions that use objects of class `"regdata"` typically require more columns. Six or seven columns (4 or 5 L-moments) is usually adequate for regional frequency analysis.

Note that the fourth column should contain values of the L-CV t, not the L-scale ℓ_2!

Objects of class `"regdata"` are created by `as.regdata`, and by `regsamlmu` (with default settings of its arguments). They are used by several functions in package **lmomRFA**, including `regavlmom` (which computes regional average L-moments), `regfit` (which fits a regional frequency distribution), and `regtst` (which computes discordancy, heterogeneity and goodness-of-fit measures).

Usage

```
as.regdata(x, warn.names=TRUE)
```

Arguments

x   R object.

warn.names   Logical: if TRUE, warnings are issued if the column names of x appear to be inconsistent with what is expected for an object of class `"regdata"`.

Details

`as.regdata` converts an R object to class `"regdata"`. Only data frames and numeric matrices can be converted.

Value

An object of class `"regdata"`.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

Examples

```
Cascades        # 'Cascades' is of class "regdata"

# Create a data frame with site statistics
dd<-data.frame(
  name      =c("site 1", "site 2", "site 3"),
  n         =c(  20,   30,   40),
  mean      =c( 100,  110,  120),
  LCV       =c(0.20, 0.25, 0.30),
  L_skewness=c(0.15, 0.20, 0.25),
  L_kurtosis=c(0.10, 0.15, 0.20),
  t_5       =c(0.10, 0.12, 0.14))
# Convert to class "regdata"
rdd<-as.regdata(dd)
rdd
class(rdd)
```

---

regfit	*Fit a regional frequency distribution*

Description

Fits a frequency distribution to a vector of regional average L-moments. Returns an object of class "rfd", which contains the specification of the regional frequency distribution: the quantile function, parameters of the regional growth curve, and the index-flood values (site-specific scale factors) for each site.

Usage

```
regfit(regdata, dist)
```

Arguments

regdata   Object of class regdata, containing summary statistics of the data for the sites in a region.

dist   Character string specifying the distribution to be fitted. See "Details" below.

Details

The function computes regional average L-moments (by calling regavlmom) and fits a probability distribution to the regional average L-moments. This distribution has mean 1, i.e., the index flood is the mean of the distribution.

For distribution dist there should exist a function to estimate the parameters of the distribution given a set of L-moments. The function should have a name that is the character string "pel" followed by the character string dist. It should accept a single argument, a vector containing L-moments ℓ_1, ℓ_2, t_3, t_4, etc., and return a vector of distribution parameters.

For distribution dist there should also exist a quantile function, which should have a name that is the character string "qua" followed by the character string dist. It should accept two arguments: a vector of probabilities and a vector containing the parameters of the distribution.

The search path used to find the ″pel″ and ″qua″ functions is the same as for arguments supplied to regfit, i.e. the enclosing frames of the function, followed by the search path specified by search().

The estimation routines and quantile functions in package **lmom** have the form described here. For example, to use a generalized extreme value distribution set dist to be the string ″gev″; then the fitting function pelgev and the quantile function quagev will be used (unless these functions have been masked by another object on the search path).

Value

An object of class ″rfd″, containing the specification of the regional frequency distribution: It is a list with the following elements:

dist	The character string dist.
para	Vector containing the parameters of the fitted regional distribution.
qfunc	The quantile function of distribution dist. It is a function that takes a single argument, a vector of probabilities, and returns a vector of quantiles.
rmom	The regional average L-moments.
index	Index flood values at each site. This is a named vector whose values are the index flood values at each site, from regdata[[3]], and whose names are the site names, from regdata[[1]].

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

Examples

```
data(Cascades)               # An object of class "regdata"
rfit <- regfit(Cascades, "gno") # Fit a generalized normal distribution
rfit                         # Print details of the fitted distribution
                             #   (components 'dist' and 'para')
rfit$index                   # Index flood values

evplot(rfit)                 # Plot the regional growth curve
evplot(qfunc=rfit$qfunc)     # The same, but with more typing
evplot(qfunc=regqfunc(rfit)) # The same, with still more typing
```

---

reglmr	*Regional weighted average of L-moments*

Description

Regional weighted average of L-moments.

Usage

```
reglmr(xmom, weight)
```

Arguments

xmom	Matrix or data frame each of whose rows contains the L-moments and L-moment ratios for one site, in the order ℓ_1, ℓ_2, t_3, t_4, etc.
weight	Vector containing the weights to be used for each site. If omitted, equal weights will be used.

Value

Vector containing the regional average L-moments.

Note

This function is deprecated and may be removed from a future version of the package. Function `regavlmom` is the recommended replacement.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

Examples

```
(xmom<-t(sapply(Maxwind,samlmu)))
nrec<-sapply(Maxwind,length)
reglmr(xmom,nrec)   # weighted by record length
reglmr(xmom)        # unweighted
```

---

regquant	*Quantiles and quantile function of a regional frequency distribution*

Description

`regquant` computes quantiles of a regional frequency distribution, i.e., values of the regional growth curve.

`regqfunc` returns a function that will compute the quantiles.

Usage

```
regquant(f, rfd)

regqfunc(rfd)
```

Arguments

f	Vector of probabilities.
rfd	Object of class `rfd`, containing the specification of a regional frequency distribution.

Value

`regquant` returns a vector of quantiles.

`regqfunc` returns the qfunc element of `rfd`. This is a function that takes one argument, which should be a vector of probabilities, and returns a vector of quantiles.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

See Also

regfit.

Examples

```
rfit <- regfit(Cascades,"gno")  # Fit regional distribution

# Compute some quantiles
regquant(seq(0.1, 0.9, by=0.1), regfit(Cascades,"gno"))

# Get the quantile function (regional growth curve)
rgc <- regqfunc(rfit)

# Compute quantiles by evaluating the regional growth curve
rgc(seq(0.1, 0.9, by=0.1))

# Plot the regional growth curve
curve(rgc, 0.01, 0.99)
```

---

regquantbounds   *Compute error bounds for a regional frequency distribution*

Description

For a regional frequency distribution, the functions compute the root mean square error (RMSE) and error bounds for quantiles either of the regional growth curve (regquantbounds) or of distributions at individual sites (sitequantbounds).

Usage

```
regquantbounds(relbounds, rfd)

sitequantbounds(relbounds, rfd, sitenames, index, seindex, drop = TRUE)
```

Arguments

relbounds	An object of class "regsimq", the result of calling function regsimq to simulate relative RMSE and error bounds for a regional frequency distribution.
rfd	An object of class "rfd", containing the specification of a regional frequency distribution.
sitenames	Vector of site names.
index	Values of the estimated site-specific scale factor ("index flood") for the sites.
seindex	Standard errors of the estimates in index.
drop	Logical: if TRUE and there is only one site, the value returned from sitequantbounds will be an object of class "rfdbounds" rather than a list containing one such object.

Details

The relative RMSE values from `relbounds` are multiplied by the quantile values from `rfd` to yield absolute RMSE values for quantile estimates, and the quantile values from `rfd` are divided by the error bounds from `relbounds` to yield error bounds for quantiles, as in Hosking and Wallis (1997), eq. (6.19). These computations apply to quantiles either of the regional growth curve (for `regquantbounds`) or of the frequency distributions at individual sites (for `sitequantbounds`).

If argument `index` of `sitequantbounds` is missing, then results (RMSE and error bounds of quantiles) are computed for sites in the region specified by `rfd` and its `index` component, assuming that the site-specific scale factor ("index flood") is estimated by the sample mean at each site, computed from the same data set that was used to fit the regional frequency distribution.

If `index` and `sitenames` are both missing, then results will be computed for all of the sites in the region specified by `rfd`.

If `index` is missing and `sitenames` is present, then error bounds will be computed for a subset of the sites in the region specified by `rfd`. `sitenames` will be used to select sites from the vector `rfd$index`, either by position or by name.

If argument `index` of `sitequantbounds` is present, then results are computed for arbitrary sites (for example, ungauged sites for which the regional growth curve of the regional frequency distribution `rfd` is believed to apply), assuming that the site-specific scale factor ("index flood") is estimated from data that are (approximately) statistically independent of the data used to fit the regional frequency distribution. In this case `relbounds$sim.rgcratio` must not be NULL, i.e. `relbounds` should have been generated by a call to `regsimq` with argument `save=TRUE`.

If `index` and `sitenames` are both present, they must have the same length, and will be taken to refer to sites whose names are the elements of `sitenames` and whose index-flood values are the elements of `index`.

If `index` is present and `sitenames` is missing, results are computed for sites whose index-flood values are the elements of `index`; if `index` has names, these names will be used as the site names.

When `index` and `seindex` are specified, it is assumed in the simulation procedure that the relative estimation error is lognormally distributed, i.e. that the logarithm of the ratio of the estimated to the true index flood value has a normal distribution with mean 0 and standard deviation `seindex/index`.

As noted by Hosking and Wallis (1997, discussion following (6.19)), error bounds in the lower tail of the distribution may be unhelpful when the fitted distribution can take negative values. In these cases the computed bounds will be NA (if the quantile estimate is negative) or Inf (if the quantile estimate is positive but the corresponding error bound in `relbounds` is negative).

Value

For `regquantbounds`, an object of class `"rfdbounds"`. This is a data frame with columns f, probabilities for which quantiles are estimated; qhat, estimated quantiles; RMSE, RMSE of the estimated quantiles. Also, for each bound probability in `relbounds$boundprob`, there is a column containing the error bound corresponding to that probability. The object also has an attribute `"boundprob"` that contains the bound probabilities.

For `sitequantbounds`, a list each of whose components is an object of class `"rfdbounds"` containing results for one site. In this case the second column of the data frame is named Qhat, not qhat. If `drop` is TRUE and the list has one component, a single `"rfdbounds"` object is returned.

Note

For a region that is confidently believed to be homogeneous, the region used to generate the results in `relbounds` may be the same as that specified by `rfd`. In practice, it is often acknowledged

that some degree of heterogeneity is present in the data to which the distribution rfd is fitted. The simulations used in function regsimq to generate relbounds can then be based on a region whose specification includes an appropriate degree of heterogeneity, and the error bounds calculated by regquantbounds and sitequantbounds will honestly reflect the failure of the assumption of homogeneity made by regfit (i.e. that the at-site growth curves are the same for all sites in the region) to hold exactly. The example below illustrates this practice.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

regsimq, which runs the simulations that generate the results returned by regquantbounds.

Examples

```
data(Cascades)              # A regional data set

rmom <- regavlmom(Cascades) # Regional average L-moments

# Fit a generalized normal distribution to the regional data
rfit <- regfit(Cascades, "gno")

# Set up an artificial region to be simulated:
# -- Same number of sites as Cascades
# -- Same record lengths as Cascades
# -- Same site means as Cascades
# -- L-CV varies linearly across sites, with mean value equal
#    to the regional average L-CV for the Cascades data.
#     'LCVrange' specifies the  range of L-CV across the sites,
#    and is chosen to reflect the amount of heterogeneity that
#    may reasonably be believed to be present in the Cascades
#    data (see the example for 'regsimh').
# -- L-skewness is the same at each site, and is equal to the
#    regional average L-skewness for the Cascades data
nsites <- nrow(Cascades)
means <- Cascades$mean
LCVrange <- 0.025
LCVs <- seq(rmom[2]-LCVrange/2, rmom[2]+LCVrange/2, len=nsites)
Lskews<-rep(rmom[3], nsites)

# Each site will have a generalized normal distribution:
# get the parameter values for each site
pp <- t(apply(cbind(means, means*LCVs ,Lskews), 1, pelgno))
pp

# Set correlation between each pair of sites to 0.64, the
# average inter-site correlation for the Cascades data
avcor <- 0.64
```

```
# Run the simulation.  To save time, use only 100 replications.
simq <- regsimq(qfunc=quagno, para=pp, cor=avcor, nrec=Cascades$n,  nrep=100, fit="gno")

# Apply the simulated bounds to the estimated regional growth curve
regquantbounds(simq, rfit)

# Apply the simulated bounds to quantiles for site 3
sitequantbounds(simq, rfit, site=3)

# Apply the simulated bounds to quantiles for a site whose mean
# is estimated to be 100 with standard error 25
sitequantbounds(simq, rfit, index=100, seindex=25)
```

---

regsamlmu	*Sample L-moments of multiple data sets*

Description

Computes the "unbiased" sample L-moments and L-moment ratios of multiple sets of data stored in a list or matrix. Following the paradigm of regional frequency analysis, we regard the data sets as coming from different measurement sites.

Usage

```
regsamlmu(x, nmom = 5, sort.data = TRUE, lcv = TRUE)
```

Arguments

x	A list of numeric vectors, or a numeric matrix.
nmom	Number of L-moments to be computed.
sort.data	Logical: whether each data set should be sorted.
lcv	Logical. If TRUE, the second L-moment will be expressed as a fraction of the mean, i.e. the computed value will be the sample L-CV $t = \ell_2/\ell_1$. If FALSE, the second L-moment will simply be the sample L-scale value ℓ_2.

Details

Sample L-moments are computed for each data set. The calculations use samlmu internally. If x is a list, each list element should contain data for one site and the names of the list elements should be the site names. If x is a matrix, each column should contain data for one site and the column names should be the site names.

Value

An object of class regdata. It is a data frame with columns "name" and "n", containing respectively the site names and the number of non-missing data values at each site, and further columns containing the L-moments and L-moment ratios, in the order ℓ_1, t (or ℓ_2), t_3, t_4, etc.

Note

The default parameter values are chosen to be convenient for the regional frequency analysis methods described by Hosking and Wallis (1997). Note that the number of L-moments and the choice of whether to return L-CV or L-scale are different from the defaults for `samlmu`.

Users of the LMOMENTS Fortran package, version 3.04, should note that its `PROGRAM XFIT` by default uses plotting-position estimators of L-moment ratios, which give different results from the "unbiased" estimators used by `regsamlmu` (and by all other functions in package **lmomRFA**).

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

Examples

```
data(Maxwind)       # a list
regsamlmu(Maxwind)

data(airquality)    # a data frame
regsamlmu(airquality[1:4])
```

---

regsimh   *Simulate the distribution of heterogeneity and goodness-of-fit measures*

Description

Estimates, using Monte Carlo simulation, the distribution of heterogeneity and goodness-of-fit measures for regional frequency analysis. These are the statistics H and Z^{DIST} defined respectively in sections 4.3.3 and 5.2.3 of Hosking and Wallis (1997).

Usage

```
regsimh(qfunc, para, cor = 0, nrec, nrep = 500, nsim = 500)
```

Arguments

qfunc	List containing the quantile functions for each site. Can also be a single quantile function, which will be used for each site.
para	Parameters of the quantile functions at each site. If qfunc is a list, para must be a list of the same length whose components are numeric vectors, the parameters of the corresponding component of qfunc. If qfunc is a single quantile function, para can be a single vector, containing a single set of parameter values that will be used for each site; a matrix or data frame whose rows each contain the parameter values for one site; or a list of length length(nrec) whose components are numeric vectors, each containing the parameter values for one site.

cor	Specifies the correlation matrix of the frequency distribution of each site's data. Can be a matrix (which will be rescaled to a correlation matrix if necessary) or a constant (which will be taken as the correlation between each pair of sites).
nrec	Numeric vector containing the record lengths at each site.
nrep	Number of simulated regions.
nsim	Number of simulations used, within each of the nrep simulated regions, when calculating heterogeneity and goodness-of-fit measures.

Details

A realization is generated of data simulated from the region specified by parameters qfunc, para, and cor, and with record lengths at each site specified by argument nrec. The simulation procedure is as described in Hosking and Wallis (1997), Table 6.1, through step 3.1.2. Heterogeneity and goodness-of-fit measures are computed for the realization, using the same method as in function regtst. The entire procedure is repeated nrep times, and the values of the heterogeneity and goodness-of-fit measures are saved. Average values, across all nrep realizations, of the heterogeneity and goodness-of-fit measures are computed.

Value

An object of class "regsimh". This is a list with the following components:

nrep	The number of simulated regions (argument nrep).
nsim	The number of simulation used within each region (argument nsim).
results	Matrix of dimension $8 \times$ nrep, containing the values, for each of the nrep simulated regions, of the heterogeneity and goodness-of-fit measures.
means	Vector of length 8, containing the mean values, across the nrep simulated regions, of the three heterogeneity and five goodness-of-fit measures.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

regtst for details of the heterogeneity and goodness-of-fit measures.

Examples

```
## Not run:
data(Cascades)           # A regional data set

rmom<-regavlmom(Cascades) # Regional average L-moments

# Set up an artificial region to be simulated:
# -- Same number of sites as Cascades
# -- Same record lengths as Cascades
# -- Mean 1 at every site (results do not depend on the site means)
```

```
# -- L-CV varies linearly across sites, with mean value equal
#    to the regional average L-CV for the Cascades data.
#    'LCVrange' specifies the  range of L-CV across the sites.
# -- L-skewness the same at each site, and equal to the regional
#    average L-skewness for the Cascades data
nsites <- nrow(Cascades)
means <- rep(1,nsites)
LCVrange <- 0.025
LCVs <- seq(rmom[2]-LCVrange/2, rmom[2]+LCVrange/2, len=nsites)
Lskews<-rep(rmom[3], nsites)

# Each site will have a generalized normal distribution:
# get the parameter values for each site
pp <- t(apply(cbind(means, means*LCVs ,Lskews), 1, pelgno))

# Set correlation between each pair of sites to 0.64, the
# average inter-site correlation for the Cascades data
avcor <- 0.64

# Run the simulation.  It will take some time (about 25 sec
# on a Lenovo W500, a moderately fast 2011-vintage laptop)
# Note that the results are consistent with the statement
# "the average H value of simulated regions is 1.08"
# in Hosking and Wallis (1997, p.98).
set.seed(123)
regsimh(qfunc=quagno, para=pp, cor=avcor, nrec=Cascades$n,
  nrep=100)

## End(Not run)
```

| regsimq | *Compute error bounds for a fitted regional frequency distribution* |

Description

Computes, using Monte Carlo simulation, relative error bounds for estimated quantiles of a regional frequency distribution fitted by `regfit`.

Usage

```
regsimq(qfunc, para, cor = 0, index = NULL, nrec, nrep = 10000,
        fit = "gev", f = c(0.01, 0.1, 0.5, 0.9, 0.99, 0.999),
        boundprob = c(0.05, 0.95), save = TRUE)
```

Arguments

qfunc	List containing the quantile functions for each site. Can also be a single quantile function, which will be used for each site.
para	Parameters of the quantile functions at each site. If qfunc is a list, para must be a list of the same length whose components are numeric vectors, the parameters of the corresponding component of qfunc. If qfunc is a single quantile function, para can be a single vector, containing a single set of parameter values that will

be used for each site; a matrix or data frame whose rows each contain the parameter values for one site; or a list of length length(nrec) whose components are numeric vectors, each containing the parameter values for one site.

cor   Specifies the correlation matrix of the frequency distribution of each site's data. Can be a matrix (which will be rescaled to a correlation matrix if necessary) or a constant (which will be taken as the correlation between each pair of sites).

index   Specifies the value of the site-specific scale factor ("index flood") at each site. Can be: a vector, containing the values at each site; a constant, which will be taken to be the index flood value at each site; or (the default) NULL, in which case the index flood will be computed by numerical integration of the quantile function.

nrec   Numeric vector containing the record lengths at each site.

nrep   Number of simulated regions.

fit   Character string specifying the distribution to be fitted. See "Details" below.

f   Vector of probabilities corresponding to the quantiles whose accuracy is to be estimated.

boundprob   Vector of probabilities for which error bounds will be computed.

save   Logical. Should the simulated values of the ratio of the estimated to the true regional growth curve be saved? These values are needed when sitequantbounds is called with its argument index present, e.g. to compute error bounds for quantiles at sites other than those whose data were used to fit the regional frequency distribution (e.g., ungauged sites). If this computation is not required, storage can be saved by setting save to FALSE.

Details

A realization of data from a region is generated as follows. The frequency distributions at the sites (specified by arguments qfunc and para) are randomly permuted, and data are simulated from these frequency distributions with inter-site correlation specified by argument cor and record lengths at each site specified by argument nrec. The regional frequency distribution specified by argument fit is then fitted to the simulated data, as in function regfit. The procedure is as described in Hosking and Wallis (1997), Table 6.1, except that the permutation of frequency distributions is a later modification, intended to give more realistic sets of simulated data.

From each realization the sample mean values and estimates of the quantiles of the regional growth curve, for nonexceedance probabilities specified by argument f, are saved.

From the simulated values, for each quantile specified by argument f the relative root mean square error (relative RMSE) is computed as in Hosking and Wallis (1997, eq. (6.15)). Error bounds are also computed, as in Hosking and Wallis (1997, eq. (6.18)) but with bound probabilities specified by argument boundprob rather than the fixed values 0.05 and 0.95 considered by Hosking and Wallis. The error bounds are sample quantiles, across the nrep realizations, of the ratio of the estimated regional growth curve to the true at-site growth curve or of the ratio of the estimated to the true quantiles at individual sites.

For distribution fit there should exist a function to estimate the parameters of the distribution given a set of L-moments. The function should have a name that is the character string "pel" followed by the character string fit. It should accept a single argument, a vector containing L-moments ℓ_1, ℓ_2, t_3, t_4, etc., and return a vector of distribution parameters.

For distribution fit there should also exist a quantile function, which should have a name that is the character string "qua" followed by the character string fit. It should accept two arguments: a vector of probabilities and a vector containing the parameters of the distribution.

The search path used to find the $"pel"$ and $"qua"$ functions is the same as for arguments supplied to regsimq, i.e. the enclosing frames of the function, followed by the search path specified by search().

The estimation routines and quantile functions in package **lmom** have the form described here. For example, to use a generalized extreme value distribution set fit to be the string $"gev"$; then the fitting function pelgev and the quantile function quagev will be used (unless these functions have been masked by another object on the search path).

Value

An object of class $"regsimq"$. This is a list with the following components:

f   Vector of probabilities corresponding to the quantiles whose accuracy is to be estimated. A copy of argument f.

boundprob   Vector of probabilities corresponding to the error bounds. A copy of argument boundprob.

relbounds.rgc   Data frame containing the relative RMSE and error bounds for the regional growth curve. It has columns $"f"$, probabilities corresponding to each quantile, $"rel.RMSE"$, relative RMSE of quantiles of regional growth curve, and, for each bound probability in boundprob, a column giving the error bound (quantile of the empirical distribution of simulated values of the ratio of the estimated regional growth curve to the true at-site growth curve) for that bound probability.

relbounds.by.site
                        List of length(nrec) data frames. Each data frame contains the relative RMSE and error bounds for quantiles at one site, and has the same structure as the relbounds.rgc component of the return value.

sim.rgcratio   If save is TRUE, a matrix of dimension length(f) $\times$ nrep, containing the simulated values of the ratio of the estimated to the true regional growth curve for quantiles corresponding to probabilities in f.

                        If save is FALSE, list element sim.rgcratio is NULL.

Note

Error bounds for the regional growth curve apply to the regional growth curve regarded as an estimator of the growth curve for a randomly chosen site. The growth curve for site i is defined to be $q_i(\,.\,) = Q_i(\,.\,)/\mu_i$ where $Q_i(\,.\,)$ is the site's quantile function and μ_i is the site-specific scale factor ("index flood") for site i. For each of the nrep simulated regions, and each probability F in f, the regional growth curve $\hat{q}(F)$ is estimated and the ratios $\hat{q}(F)/q_i(F)$ are calculated. The relative error bounds are empirical quantiles, corresponding to the probabilities in boundprob, of the nrep * length(nrec) values of these ratios obtained from the simulations.

This differs from the approach taken in Hosking and Wallis (1997), Table 6.2 and Fig. 6.2, in which error bounds are computed regarding the estimated regional growth curve as an estimator of the regional average growth curve $q^{\mathrm{R}}(\,.\,)$, the harmonic mean of the sites' growth curves (Hosking and Wallis, 1997, p. 102).

When the parent region is homogeneous, with identical frequency distributions at each site (apart from a site-specific scale factor), the two approaches give identical results. For heterogeneous regions the "regard as estimator of randomly chosen site growth curve" approach yields error bounds that are wider, but are more realistic given that the ultimate aim of regional frequency analysis is estimation of quantiles at individual sites.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

regfit, for details of fitting a regional frequency distribution; regquantbounds and sitequantbounds, for converting the relative bounds returned by regsimq into absolute bounds for quantiles of the regional growth curve or of the frequency distributions at individual sites.

Examples

```
data(Cascades)              # A regional data set

rmom <- regavlmom(Cascades) # Regional average L-moments

# Fit generalized normal distribution to regional data
rfit <- regfit(Cascades, "gno")

# Set up an artificial region to be simulated:
# -- Same number of sites as Cascades
# -- Same record lengths as Cascades
# -- Same site means as Cascades
# -- L-CV varies linearly across sites, with mean value equal
#    to the regional average L-CV for the Cascades data.
#    'LCVrange' specifies the  range of L-CV across the sites.
# -- L-skewness the same at each site, and equal to the regional
#    average L-skewness for the Cascades data
nsites <- nrow(Cascades)
means <- Cascades$mean
LCVrange <- 0.025
LCVs <- seq(rmom[2]-LCVrange/2, rmom[2]+LCVrange/2, len=nsites)
Lskews<-rep(rmom[3], nsites)

# Each site will have a generalized normal distribution:
# get the parameter values for each site
pp <- t(apply(cbind(means, means*LCVs ,Lskews), 1, pelgno))

# Set correlation between each pair of sites to 0.64, the
# average inter-site correlation for the Cascades data
avcor <- 0.64

# Run the simulation.  To save time, use only 100 replications.
simq <- regsimq(qfunc=quagno, para=pp, cor=avcor, nrec=Cascades$n,
  nrep=100, fit="gno")

# Relative RMSE and error bounds for the regional growth curve
simq$relbounds.rgc

# Relative RMSE and error bounds for quantiles at site 3
simq$relbounds.by.site[[3]]
```

---

regtst   *Test statistics for regional frequency analysis*

Description

Computes discordancy, heterogeneity and goodness-of-fit measures for regional frequency analysis. These are the statistics D_i, H, and Z^{DIST} defined respectively in sections 3.2.3, 4.3.3, and 5.2.3 of Hosking and Wallis (1997).

Usage

```
regtst(regdata, nsim=1000)

regtst.s(regdata, nsim=1000)
```

Arguments

regdata   Object of class [regdata](#) containing the input data. It should be a data frame, each of whose rows contains data for one site. The first seven columns should contain respectively the site name, record length and L-moments and L-moment ratios, in the order ℓ_1 (mean), t (L-CV), t_3 (L-skewness), t_4 (L-kurtosis), and t_5.

Note that the fourth column should contain values of the L-CV t, not the L-scale ℓ_2!

Function regsamlmu, with default settings of its arguments, returns an object of class "regdata".

nsim   Number of simulations to use in the calculation of the heterogeneity and goodness-of-fit measures.

If less than 2, only the discordancy measure will be calculated.

Details

The discordancy measure D_i indicates, for site i, the discordancy between the site's L-moment ratios and the (unweighted) regional average L-moment ratios. Large values might be used as a flag to indicate potential errors in the data at the site. "Large" might be 3 for regions with 15 or more sites, but less (exact values in list element Dcrit) for smaller regions.

Three heterogeneity measures are calculated, each based on a different measure of between-site dispersion of L-moment ratios: [1] weighted standard deviation of L-CVs; [2] average of L-CV/L-skew distances; [3] average of L-skew/L-kurtosis distances. These dispersion measures are the quantities V, V_2, and V_3 defined respectively in equations (4.4), (4.6), and (4.7) of Hosking and Wallis (1997). The heterogeneity measures are calculated from them as in equation (4.5) of Hosking and Wallis (1997). In practice H[1] is probably sufficient. A value greater than (say) 1.0 suggests that further subdivision of the region should be considered as it might improve the accuracy of quantile estimates.

Goodness of fit is evaluated for five candidate distributions: generalized logistic, generalized extreme value, generalized normal (lognormal), Pearson type III (3-parameter gamma), and generalized Pareto. In the output the distributions are referred to by 3-letter abbreviations, respectively glo, gev, gno, pe3, and gpa. If the region is homogeneous and data at different sites are statistically independent, then if one of the distributions is the true distribution for the region its goodness-of-fit measure should have approximately a standard normal distribution. Provided that the region is

acceptably close to homogeneous, the fit may be judged acceptable at the 10 per cent significance level if the Z value is less than 1.645 (i.e., qnorm(0.95)) in absolute value.

Calculation of heterogeneity and goodness-of-fit measures involves the sampling variability of L-moment ratios in a homogeneous region whose record lengths and average L-moment ratios match those of the data. The sampling variability is estimated by Monte Carlo simulation using nsim replications of the region. Results will vary between invocations of regtst with different seeds for the random-number generator.

In the homogeneous region used in the simulations, the sites have a kappa distribution, fitted to the regional average L-moment ratios of the data in regdata. The kappa fit may fail if the regional average L-kurtosis is high relative to the regional average L-skewness. In this case a kappa distribution is fitted with shape parameter h constrained to be -1 (i.e., a generalized logistic distribution); this gives the largest possible L-kurtosis value for a kappa distribution with given L-skewness.

regtst and regtst.s are functionally identical. regtst calls a Fortran routine internally and is faster, typically by a factor of 3 or 4. regtst.s is written almost entirely in the S language; it is provided so that users can see how the calculations are done, and can conveniently alter the code for their own purposes if necessary.

Value

An object of class "regtst", which is a list with elements as follows.

data	The input data, i.e. data frame regdata after coercion to class "regdata" if necessary.
nsim	Number of simulations, i.e. the argument nsim.
D	Vector containing the discordancy measures for each site.
Dcrit	Vector of length 2 containing critical values of the discordancy measure corresponding to significance levels of 10 and 5 per cent — except that the values never exceed 3 and 4 respectively. See Hosking and Wallis (1997), section 3.2.4.
rmom	Vector of length 5 containing the regional weighted average L-moment ratios (weights proportional to record lengths).
rpara	Vector of length 4 containing the parameters of a kappa distribution fitted to the regional weighted average L-moment ratios.
vobs	Vector of length 3 containing the observed values of the three measures of between-site dispersion of L-moment ratios.
vbar	Vector of length 3 containing the mean of the simulated values of the three dispersion measures.
vsd	Vector of length 3 containing the standard deviation of the simulated values of the three dispersion measures.
H	Vector of length 3 containing the three measures of regional heterogeneity.
para	List of length 6 containing the parameters of the five candidate distributions and the Wakeby distribution (3-letter abbreviation "wak") fitted to the regional weighted average L-moment ratios.
t4fit	Vector of length 5 containing the L-kurtosis of the five candidate distributions fitted to the regional weighted average L-moment ratios.
Z	Vector of length 5 containing the goodness-of-fit measures for each of the five candidate distributions.

Note

Data frame `regdata` may have only six columns, i.e. the fifth L-moment ratio t_5 may be omitted. In this case the return value will contain missing values for `rmom[5]` and the elements of `para$wak`.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M. (1996). Fortran routines for use with the method of L-moments, Version 3. Research Report RC20525, IBM Research Division, Yorktown Heights, N.Y.

Hosking, J. R. M., and Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press.

See Also

`summary.regtst` for summaries.

Examples

```
# An example from Hosking (1996).  Compare the output with
# the file 'cascades.out' in the LMOMENTS Fortran package at
# http://lib.stat.cmu.edu/lmoments/general (results will not
# be identical, because random-number generators are different).
summary(regtst(Cascades, nsim=500))

# Output from 'regsamlmu' can be fed straight into 'regtst'
regtst(regsamlmu(Maxwind))
```

---

sitequant	*Quantiles and quantile functions for individual sites in a region*

Description

Quantiles and quantile functions for individual sites in a region. `sitequant` computes quantiles directly; `siteqfunc` returns a function that will compute quantiles.

Usage

```
sitequant(f, rfd, sitenames, index, drop = TRUE)

siteqfunc(rfd, sitenames, index)
```

Arguments

f	Vector of probabilities.
rfd	Object of class [rfd](), containing the specification of a regional frequency distribution.
sitenames	Vector of site names.
index	Values of the site-specific scale factor ("index flood") for the sites.
drop	Logical: if TRUE and there is only one site, or one probability value, the value returned from `sitequant` will be a vector rather than a matrix.

Details

If index and sitenames are both present, they must have the same length, and will be taken to refer to sites whose names are the elements of sitename and whose index-flood values are the elements of index.

If index is present and sitenames is missing, quantiles are computed for sites whose index-flood values are the elements of index; if index has names, these names will be used as the site names.

If sitenames is present and index is missing, then quantiles will be computed for a subset of the sites in the region specified by rfd. sitenames will be used to select sites from the vector rfd$index, either by position or by name.

If sitenames and index are both missing, then quantiles will be computed for all of the sites in the region specified by rfd.

Value

For sitequant, a matrix whose rows contain quantiles for a single site, for the probabilities specified in f. If drop is TRUE and the matrix has only one row or column, it will be returned as a vector.

For siteqfunc, a function or a list of functions that each compute quantiles for one site. Each function takes a single argument, a vector of probabilities, and returns a vector of quantiles.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

Examples

```
rfit <- regfit(Cascades, "gno")   # Fit regional distribution

## Quantiles for:
# - sites in the Cascades data set, indexed by number
sitequant(c(0.9, 0.99, 0.999), rfit, sitenames=1:3)

# - sites in the Cascades data set, indexed by name
sitequant(c(0.9, 0.99, 0.999), rfit,
  sitenames=c("350304", "351433", "351862"))

# - other sites, with specified index floods
sitequant(c(0.9, 0.99, 0.999), rfit, index=c(80, 100))

# - other sites, with specified index floods and names
sitequant(c(0.9, 0.99, 0.999), rfit, index=c(80, 100),
  sitenames=c("Site 80", "Site 100"))

# - a single site, with drop=FALSE: result is a matrix
sitequant(c(0.9, 0.99, 0.999), rfit, sitenames=10, drop=FALSE)

# - a single site, with drop=TRUE (the default): result is a vector
sitequant(c(0.9, 0.99, 0.999), rfit, sitenames=10)

# Quantile function for site 10
qfunc10 <- siteqfunc(rfit, site=10)

# Compute quantiles via the quantile function
qfunc10(c(0.9, 0.99, 0.999))
```

```
# Plot the quantile function
evplot(qfunc=qfunc10)
```

---

summary.regtst   *Summary of test statistics for regional frequency analysis*

Description

summary method for an object of class "regtst".

Usage

```
## S3 method for class 'regtst'
summary(object,
  prob = c(0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.8, 0.9, 0.95, 0.98, 0.99, 0.999),
  conf = 0.90, decimals = c(4, 4, 2, 3), ...)

## S3 method for class 'summary.regtst'
print(x, decimals, ...)
```

Arguments

object	An object of class "regtst", usually the result of a call to regtst.
x	An object of class "summary.regtst", usually the result of a call to summary.regtst.
prob	Nonexceedance probabilities for which quantile estimates should be printed.
conf	Confidence level for printing parameter and quantile estimates. These quantities will be printed only for distributions that give an adequate fit at the specified confidence level.
decimals	Vector of length 4. The four elements specify the number of decimal places to be used when printing L-moment ratios, distribution parameters, test statistics, and quantile estimates, respectively.
...	Further arguments passed to or from other methods.

Details

The printed output corresponds closely to that produced by function REGTST in the LMOMENTS Fortran package (Hosking, 1996).

Value

summary.regtst and print.summary.regtst each return, invisibly, an object of class "summary.regtst", which is a list with elements as for class "regtst", plus the following elements:

conf	Confidence level — the conf argument supplied to summary.regtst.
prob	Vector of nonexceedance probabilities — the prob argument supplied to summary.regtst.
quant	Matrix with 6 rows and length(prob) columns, containing quantile estimates for the five candidate distributions and the Wakeby distribution.
decimals	Vector of length 4. Number of decimals to be used when printing an object of class "summary.regtst" if the decimals argument of print.summary.regtst is not specified.

Author(s)

J. R. M. Hosking <jrmhosking@gmail.com>

References

Hosking, J. R. M. (1996). Fortran routines for use with the method of *L*-moments, Version 3. Research Report RC20525, IBM Research Division, Yorktown Heights, N.Y.

See Also

regtst

Examples

```
# An example from Hosking (1996).  Compare the output with
# the file 'cascades.out' in the LMOMENTS Fortran package at
# http://lib.stat.cmu.edu/lmoments/general (results will not
# be identical, because random-number generators are different).
summary(regtst(Cascades, nsim=500))
```

Index