Package ‘rich’

October 14, 2022

Type Package

Title Computes and Compares Species Richnesses

Version 1.0.1

Date 2016-12-08

Author Jean-Pierre Rossi

Maintainer Jean-Pierre Rossi <jean-pierre.rossi@supagro.inra.fr>

Depends R (>= 3.1.0)

Imports vegan, boot

Description Computes rarefaction curves, cumulated and mean species richness. Compares these estimates by means of randomization tests.

License GPL (>= 2)

LazyLoad TRUE

Suggests knitr, gplots, rmarkdown

VignetteBuilder knitr

Repository CRAN

Repository/R-Forge/Project rich

Repository/R-Forge/Revision 5

Repository/R-Forge/DateTimeStamp 2016-08-12 09:22:54

Date/Publication 2016-08-12 16:36:36

NeedsCompilation no

R topics documented:

`c2cv` *Comparison of 2 values of species richness using a randomization procedure*

Description

Species richnesses are computed as the cumulative value over all samples. Richnesses are compared by mean of a randomization test without controlling for differences of sampling regime of communities density.

Usage

```
c2cv(com1,com2,nrandom=99,pr1=0.025,pr2=0.975,verbose=TRUE)
```

Arguments

`com1`	A first species-sample matrix (community 1). Rows correspond to samples whereas columns stand for species.
`com2`	A second species-sample matrix (community 2). Rows correspond to samples whereas columns stand for species.
`nrandom`	Number of randomizations to be performed. Default fixed to 99.
`pr1`	Lower probability level for quantile computations. Default fixed to 0.025.
`pr2`	Higher probability level for quantile computations. Default fixed to 0.975.
`verbose`	If `verbose` is TRUE `c2cv` returns a vector that contains the observed and randomized differences between richnesses.

Details

If the observed richness for community 1 $\geq$ observed value for community 2, `c2cv` returns a probability `p` estimated as the number of randomizations for which the observed value for community 1 $\geq$ observed value for community 2 divided by the number of randomization + 1. Similarly, if the observed value for community 1 $\leq$ observed value for community 2, `p` corresponds to the frequency of such situation in the randomizations.

Value

If `verbose==TRUE` `c2cm` returns a data frame and a vector with the randomized values. Otherwise, only the data frame is returned.

`res`	A data frame with the outputs of the randomization test: `cv1` Observed cumulative richness for community 1. `cv2` Observed cumulative richness for community 2. `cv1-cv2` Difference between observed cumulative richness of community 1 and community 2. `p` Probability of encountering such a value for `cv1-cv2` (see details above). `quantile for pr1` Quantile value for probability level `pr1`. `quantile for pr2` Quantile value for probability level `pr2`. `randomized cv1-cv2` Mean values of randomized and the observed values. `nrandom` Number of randomizations used in the test.
`rand`	A vector of `nrandom+1` values corresponding to the observed difference of `cv1-cv2` and the randomized values. `rand` is available if `verbose == TRUE`.

Note

The observed difference between populations is included in the numerator and the denominator when computing the probability `p`. This is justified because if the null hypothesis (there is no difference between populations) is true then the observed difference between populations is just another value for the randomization distribution (Manly, 1997, p. 7).

Author(s)

Jean-Pierre Rossi, <jean-pierre.rossi@supagro.inra.fr>

References

Manly, B.F.J. (1997). Randomization and Monte Carlo methods in biology. Chapman & Hall.

See Also

`c2m`, `rich`

Examples

```
## Not run:
data(efeb)
c2cv(com1=efeb$ef,com2=efeb$eb,nrandom=100,verbose=FALSE)

## End(Not run)
```

`c2m` *Compares 2 mean values using a randomization test*

Description

Mean values of 2 populations are compared using a randomization procedure. Overlapping populations are allowed.

Usage

```
c2m(pop1, pop2,pop3=NULL,nrandom,pr1=0.025,pr2=0.975,verbose=TRUE)
```

Arguments

`pop1`	A vector with the observed values for population 1.
`pop2`	A vector with the observed values for population 2.
`pop3`	A vector with the observed values that are common to population 1 and 2.
`nrandom`	Number of randomizations to perform. Default fixed to 99.
`pr1`	Lower probability level for quantile computations. Default fixed to 0.025.
`pr2`	Higher probability level for quantile computations. Default fixed to 0.975.
`verbose`	If `verbose` is TRUE `c2m` returns a vector that contains the observed and randomized differences between mean richnesses.

Details

This randomization test compares the average value of a quantitative variable sampled in 2 populations. Details are available in Manly (1997). In some cases, populations share some observed values: for example if we compare the mean annual temperature of sites where either species A or B is present and if A and B are sympatric in some localities (see example below). Those shared values are passed to `c2m` by the argument `pop3`.
If the mean value for population 1 $\geq$ mean value for population 2, `p` is the number of randomizations for which the mean value for population 1 $\geq$ mean value for population 2 divided by the number of randomizations + 1. If the mean value for population 1 $\leq$ mean value for population 2, `p` is the number of randomizations for which the mean value for population 1 $\leq$ mean value for population 2 divided by the number of randomizations + 1. If $mv1 = mv2$ `p` in not computed (`p=NC`).

Value

`res`	A data frame showing the outputs of the randomization test: `mv1` Observed mean values over samples forming population 1. `mv2` Observed mean values over samples forming population 2. `mv1-mv2` Difference between observed mean values of population 1 and population 2. `p` Probability of encountering such a value for `mv1-mv2` (see details above). `quantile for pr1` Quantile value for probability level `pr1`.

	`quantile for pr2` Quantile value for probability level `pr2`. `randomized mv1-mv2` Mean values of randomized and the observed values. `nrandom` Number of randomizations used in the test.
`rand`	A vector of `nrandom+1` values corresponding to the observed difference of `mv1-mv2` and the randomized values. `rand` is available if `verbose == TRUE`.

Note

The observed difference between populations is included in the numerator and the denominator when computing the probability `p`. This is justified because if the null hypothesis (there is no difference between populations) is true then the observed difference between populations is just another value for the randomization distribution (Manly, 1997, p. 7).

Author(s)

Jean-Pierre Rossi, <jean-pierre.rossi@supagro.inra.fr>

References

Manly, B.F.J. (1997). Randomization and Monte Carlo methods in biology. Chapman & Hall.

See Also

`c2cv`, `rich`

Examples

```
## Not run:
# The example of mandible length of male and female
# golden jackals from Manly (1997), p.4.
males<-c(120, 107, 110, 116, 114, 111, 113, 117, 114, 112)
females<-c(110, 111, 107, 108, 110, 105, 107, 106, 111, 111)
out <- c2m(pop1=males, pop2=females, nrandom=99)
out$res
hist(out$rand)
abline(v=out$res[3,1], col=”red”)
abline(v=out$res[5,1], col=”blue”)
abline(v=out$res[6,1], col=”blue”)

# Compare simulated datasets
pop1<-rnorm(10)
pop2<-rnorm(10)
out <- c2m(pop1=pop1, pop2=pop2, nrandom=99)
out$res
hist(out$rand)
abline(v=out$res[3,1], col=”red”)
abline(v=out$res[5,1], col=”blue”)
abline(v=out$res[6,1], col=”blue”)
```

```
# Maximum temperature in a set of sites where the bark beetle Tomicus destruens
# or T. piniperda are present. Both species are present at 4 sites.
data(Tomicus)
out <- c2m(pop1=Tomicus$destruens,pop2=Tomicus$piniperda,
pop3=Tomicus$both, nrandom=99)
out$res
hist(out$rand)
abline(v=out$res[3,1], col="red")
abline(v=out$res[5,1], col="blue")
abline(v=out$res[6,1], col="blue")

# using c2m with outputs of rich
data(ef)
o1 <- rich(matrix=ef, nrandom=99, verbose=TRUE)
data(ea)
o2 <- rich(matrix=ea, nrandom=99, verbose=TRUE)
out <- c2m(pop1=o1$sumrow, pop2=o2$sumrow, nrandom=999, verbose=TRUE)
hist(out$rand)
abline(v=out$res[3,1], col="red")
abline(v=out$res[5,1], col="blue")
abline(v=out$res[6,1], col="blue")

## End(Not run)
```

`ea` *Soil macrofauna in a tropical field under traditional slash-and-burn (Elahe, French Guiana)*

Description

The `ea` data set is a data frame describing the species abundance of soil macrofauna in a set of soil samples collected in a field (aka abattis) in French Guiana. These data were collected in the viciny of the Wayana village of Elahe situated on the Tampoc river.

Usage

```
data(ea)
```

Format

`ea` is a data frame containing the observations of 45 species (columns) at 30 sampling locations (rows).

Source

Rossi, J.-P., Celini, L., Mora, P., Mathieu, J., Lapied, E., Nahmani, J., Ponge, J.-F., and Lavelle, P. (2010). Decreasing fallow duration in tropical slash-and-burn agriculture alters soil macroinvertebrate diversity: A case study in southern French Guiana. Agriculture, Ecosystems and Environment 135, 148-154.

Examples

```
## Estimating species richness using rich
data(ea)
rich(matrix=ea)
```

`ef` *Soil macrofauna in a tropical forest site (Elahe, French Guiana)*

Description

The `ef` data set is a data frame describing the species abundance of soil macrofauna in a set of soil samples collected in a secondary forest in French Guiana. These data were collected in the viciny of the Wayana village of Elahe situated on the Tampoc river.

Usage

```
data(ef)
```

Format

`ef` is a data frame containing the observations of 121 species (columns) at 30 sampling locations (rows).

Source

Rossi, J.-P., Celini, L., Mora, P., Mathieu, J., Lapied, E., Nahmani, J., Ponge, J.-F., and Lavelle, P. (2010). Decreasing fallow duration in tropical slash-and-burn agriculture alters soil macroinvertebrate diversity: A case study in southern French Guiana. Agriculture, Ecosystems and Environment 135, 148-154.

Examples

```
## Estimating species richness using rich
data(ef)
rich(matrix=ef)
```

`efea` *Soil macrofauna in a tropical forest site and cultivated plot (traditional slash-and-burn, Elahe, French Guiana)*

Description

The `efea` data set is a list of two data sets. Each one is a data frame giving the species abundance of soil macrofauna in a set of soil samples. `efeb$ef` corresponds to samples collected in a secondary forest in French Guiana while `efeb$ea` contains data collected in in a field (aka abattis) in French Guiana. These data were collected in the viciny of the Wayana village of Elahe situated on the Tampoc river.

Usage

```
data(efea)
```

Format

`efea` is a list of two data frame containing the observations of 142 species (columns) at 30 sampling locations (rows).

Source

Rossi, J.-P., Celini, L., Mora, P., Mathieu, J., Lapied, E., Nahmani, J., Ponge, J.-F., and Lavelle, P. (2010). Decreasing fallow duration in tropical slash-and-burn agriculture alters soil macroinvertebrate diversity: A case study in southern French Guiana. Agriculture, Ecosystems and Environment 135, 148-154.

Examples

```
data(efea)
rich(matrix=efea[[1]])
shared(efea)
```

`efeb` *Soil macrofauna in a tropical forest site before and after deforestation (traditional slash-and-burn, Elahe, French Guiana)*

Description

The `efeb` data set is a list of two data sets. Each one is a data frame giving the species abundance of soil macrofauna in a set of soil samples. `efeb$ef` corresponds to samples collected in a secondary forest in French Guiana while `efeb$eb` contains data collected in the same site just after the forest has been cut and burnt following the traditional Amerindian slash-and-burn system. These data were collected in the viciny of the Wayana village of Elahe situated on the Tampoc river.

Usage

```
data(efeb)
```

Format

`efeb` is a list of two data frame containing the observations of 134 species (columns) at 30 sampling locations (rows).

Source

Rossi, J.-P., Celini, L., Mora, P., Mathieu, J., Lapied, E., Nahmani, J., Ponge, J.-F., and Lavelle, P. (2010). Decreasing fallow duration in tropical slash-and-burn agriculture alters soil macroinvertebrate diversity: A case study in southern French Guiana. Agriculture, Ecosystems and Environment 135, 148-154.

Examples

```
data(efeb)
rich(matrix=efeb[[1]])
shared(efeb)
```

`rarc` *Computes rarefaction curves and statistical envelop*

Description

`rarc` computes rarefaction curves and returns a data frame with bootstrap estimates of species richness, its statistical envelop and the average number of individuals for a given sample size.

Usage

```
rarc(matrix, samplesize=NULL, nrandom=99, p1=0.975, p2=0.025, save=FALSE)
```

Arguments

`matrix`	`matrix` is a typical species-sample matrix. Rows correspond to samples whereas columns stand for species.
`samplesize`	A vector containing the different sample sizes for which the computations are required. If missing or non vector, `samplesize` is replaced by a sequence from 1 to the maximum possible sample size given the size of `matrix`.
`nrandom`	`nrandom` is the number of a randomizations used in the bootstrap estimation. Values of `nrandom` <10 or NULL are set to 99.
`p1`	The probability value used to compute the upper bound of the statistical envelop using `quantile`. Default set to 0.975.
`p2`	The probability value used to compute the lower bound of the statistical envelop using `quantile`. Default set to 0.025.
`save`	If `TRUE` the function returns the raw values used in the estimation of the quantiles.

Value

`out`	A data frame with 5 columns giving the mean species richness, the lower and upper bound of the richness corresponding to the quantile values for p1 and p2, the mean number of individuals and the sample size.
`bootstrapped.val`	Generated if `save=TRUE`. A list of length of the number of different sample sizes analyzed containing the raw values used to computed the quantiles. These values can be used for example to compute the standard error as illustrated in the examples below.

Author(s)

Jean-Pierre Rossi `<jean-pierre.rossi@supagro.inra.fr>`

See Also

`raref`, `raref2`, `rich`

Examples

```
## Not run:
data(ef)
t <- rarc(ef,samplesize=c(5,10,15,20,25), nrandom=30, p1=0.975, p2=0.025)
head(t)

# Plot the outputs
plot(t$out[,6],t$out[,1], type="b", ylim=range(c(t$out[,2],t$out[,3])),
xlab="number of sampling units", ylab="richness")
points(t$out[,6] , t$out[,2], type="l", col="red")
points(t$out[,6] , t$out[,3], type="l", col="blue")

## End(Not run)

#See the package vignette for more examples. Type: vignette("rich_introduction").
```

`raref` *Estimates the species richness for a given number of sampled individuals on the basis of the corresponding rarefaction curve*

Description

`raref` computes rarefaction curve and interpolates the species richness corresponding to a given density of individuals.

Usage

```
raref(matrix, dens, nrandom=99)
```

Arguments

`matrix`	`matrix` is a typical species-sample matrix. Rows correspond to samples whereas columns stand for species.
`dens`	A number corresponding to the density for which a species richness estimation is required. `dens` must be comprised within the range of the observed density.
`nrandom`	`nrandom` is the number of a randomizations used in the bootstrap estimation. Values of `nrandom` <10 or NULL are set to 99.

Value

`raref` returns a list with two components.

`rar`	A data frame with three columns giving the observed species richness for different sample size and corresponding density of individuals.
`Sinterp`	A vector with the term dens and the corresponding interpolated species richness.

Author(s)

Jean-Pierre Rossi `<jean-pierre.rossi@supagro.inra.fr>`

References

Gotelli, N., Colwell, R. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4, 379-391.

See Also

`rarc`, `raref2`, `rich`

Examples

```
## Not run:
data(ef)
raref(ef, dens=1100, nrandom=100)

## End(Not run)
```

`raref2` *Estimates the species richness for a given number of sampled individuals on the basis of the corresponding rarefaction curve*

Description

`raref2` computes an estimation of the species richness by thinning the data matrix so that the overall corresponding density is comprised in a fixed interval.

Usage

```
raref2(matrix, dens, tolerance, nrandom=99)
```

Arguments

`matrix`	`matrix` is a typical species-sample matrix. Rows correspond to samples whereas columns stand for species.
`dens`	A number corresponding to the density for which a species richness estimation is required.`dens` must be comprised within the range of the observed density.
`tolerance`	A number used to calculate the range of acceptable values for the density of simulated communities. Let D be largest density (com1 or com2). The acceptable interval is $D \pm tolerance \times D$.
`nrandom`	`nrandom` is the number of a randomizations used in the bootstrap estimation. Values of `nrandom` <10 or NULL are set to 99.

Value

`raref2` returns a list with two components.

`mean.boot`	The mean value of the bootstrap estimates of the species richness.
`sd.boot`	The standard deviation of the bootstrap estimates.

Author(s)

Jean-Pierre Rossi <`jean-pierre.rossi@supagro.inra.fr`>

See Also

`raref`, `rich`

Examples

```
## Not run:
data(ef)
raref2(matrix=ef,dens=1100,tolerance=0.01,nrandom=50)

## End(Not run)
```

`rich` *Species richness, their confidence interval and other useful indices*

Description

Computes the cumulative and average species richness over a set of samples, the associated bootstrap statistics and other useful indices.

Usage

```
rich(matrix, verbose = FALSE, nrandom = NULL)
```

Arguments

`matrix`	`matrix` is a typical species-sample matrix. Rows correspond to samples whereas columns stand for species.
`verbose`	If `verbose=FALSE`, a simplied output is returned.
`nrandom`	Number or randomizations if bootstrap estimations are required. Non-null values < 10 are set to 99.

Details

`rich` computes basic descriptive statistics from typical species by sample data sets.

Value

`cr`	Cumulated richness over sampling units.
`mr`	Mean richness over sampling units.
`mrsd`	Standard deviation of the mean richness.
`singletons`	Species with at most one observation.
`doubletons`	Species with at most two observations.
`uniques`	Species encountered in only one sample.
`duplicates`	Species encountered in only two samples.
`bootCR`	A data frame showing the outputs of the bootstrap analyses of the cumulative richness (computed if `nrandom != FALSE`): `cr.obs` Observed cumulative richness, equal to `cr` above. `cr.boot` Mean of the bootstrap values. `cr.bcorr` Mean of the bootstrap values corrected for the bias estimated below. `cr.bias` Bias. `cr.se` Standard error of the cumulative richness estimated by bootstrap. This is estimated as the standard deviation of the bootstrap values. `cr.lbn` Lower bound of the first order normal approximation confidence interval (see `boot.ci {boot}`). `cr.ubn` Upper bound of the first order normal approximation confidence interval (see `boot.ci {boot}`).
`bootMR`	A data frame showing the outputs of the bootstrap analyses of the mean richness (computed if `nrandom != FALSE`): `mr.obs` Observed mean richness, equal to `mr` above. `mr.boot` Mean of the bootstrap values. `mr.bcorr` Mean of the bootstrap values corrected for the bias estimated below. `mr.bias` Bias. `mr.se` Standard error of the mean richness estimated by bootstrap. This is estimated as the standard deviation of the bootstrap values.

	`mr.lbn` Lower bound of the first order normal approximation confidence interval (see `boot.ci {boot}`). `mr.ubn` Upper bound of the first order normal approximation confidence interval (see `boot.ci {boot}`).
`nrandom`	Number of randomizations used in the bootstrap.
`richvec`	A vector containing the observed richness in each sample or populations. The mean of `richvec` corresponds to `mr` (available if `verbose == TRUE`).
`matrix`	The data set passed to `rich` as the `matrix` argument (available if `verbose == TRUE`).
`matrixbin`	Binary (presence/absence) transformation of the data set passed to `rich` as the `matrix` argument (available if `verbose == TRUE`).
`sumrow`	A vector of the sum of species for each sample (available if `verbose == TRUE`).
`sumcol`	A vector of the sum of sample for each species (available if `verbose == TRUE`).
`zeroes`	Number of zeroes in the data set (available if `verbose == TRUE`).

Author(s)

Jean-Pierre Rossi `<jean-pierre.rossi@supagro.inra.fr>`

References

Manly, B.F.J. (1997). Randomization and Monte Carlo methods in biology. Chapman & Hall.

See Also

`rarc`, `raref2`, `raref`

Examples

```
data(ef)
# No bootstrap statistics
rich(matrix=ef)

## Not run:
# Bootstrap estimation based on 499 randomizations
rich(matrix=ef, nrandom=499)
## End(Not run)
```

`shared` *Computes the number of species shared by groups of sampling units and other useful statistics*

Description

`shared` computes the richness of each group of sample depicting a community, the number of species shared by pairs of communities and the total number of species for each pairs of community.

Usage

```
shared(lmatrix)
```

Arguments

`lmatrix`	A list of data matrices representing the species composition of local communities. All matrices must have the same size.

Value

A data frame with: on the diagonal the observed richness for each local community, above the diagonal: the richness common to pairs of communities (shared richness) and below the diagonal: the total richness for pooled pairs of communities.

Author(s)

Jean-Pierre Rossi `<jean-pierre.rossi@supagro.inra.fr>`

References

Empirical data have been presented in a way very similar to what is returned by `shared` in the following papers:

Rossi, J.-P., Mathieu, J., Cooper, M., and Grimaldi, M. (2006). Soil macrofaunal biodiversity in amazonian pastures: Matching sampling with patterns. Soil Biology & Biochemistry, 38, 2178-2187.

Rossi, J.-P., Celini, L., Mora, P., Mathieu, J., Lapied, E., Nahmani, J., Ponge, J.-F., and Lavelle, P. (2010). Decreasing fallow duration in tropical slash-and-burn agriculture alters soil macroinvertebrate diversity: A case study in southern French Guiana. Agriculture, Ecosystems and Environment 135, 148-154.

See Also

`rich`

Examples

```
data(efeb)
shared(efeb)

sp1<-c(1,2,3,4,5)
sp2<-c(0,0,0,0,0)
sp3<-c(1,1,0,0,0)
sp4<-c(0,0,0,0,0)
site1<-cbind(sp1, sp2, sp3, sp4)
colnames(site1)<-c("sp1", "sp2", "sp3", "sp4")
sp1<-c(1,2,3)
sp2<-c(1,0,0)
sp3<-c(0,0,0)
sp4<-c(0,0,0)
site2<-cbind(sp1, sp2, sp3, sp4)
colnames(site2)<-c("sp1", "sp2", "sp3", "sp4")
sp1<-c(1,2,3,4)
sp2<-c(1,0,0,0)
sp3<-c(1,0,0,0)
sp4<-c(1,0,0,0)
site3<-cbind(sp1, sp2, sp3, sp4)
colnames(site3)<-c("sp1", "sp2", "sp3", "sp4")

data<-list(site1,site2, site3)
names(data)<-c("site1","site2","site3")
shared(data)
```

`Tomicus` *Maximum temperature at sites hosting 2 species of closely related bark-beetles of genus Tomicus*

Description

The `Tomicus` data set consists of a list with 3 components:
`destruens` Maximum temperature at 33 sites where *T. destruens* was observed.
`piniperda` Maximum temperature at 20 sites where *T. piniperda* was observed.
`both` Maximum temperature at 4 sites where both *T. destruens* and *T. piniperda* were observed in sympatry.

Usage

```
data(Tomicus)
```

Format

`Tomicus` is a list of 3 vectors.

Source

Horn, A. (2006) Comparaison des distributions passee et presente de deux especes proches de scolytes, *Tomicus piniperda* et *T. Destruens* (Coleoptera: Scolytinae). These de l'Universite d'Orleans, France 180pp.

Examples

```
## Comparing mean temperatures at sites where either T. destruens,
## T. piniperda or both are present using comp2mean
data(Tomicus)
c2m(pop1=Tomicus$destruens,pop2=Tomicus$piniperda,
pop3=Tomicus$both, nrandom=99)
```

Index