Package ‘Rarefy’

March 12, 2021

Type Package

Title Rarefaction Methods

Version 1.1

Date 2021-03-11

Description Includes functions for the calculation of spatially and non-spatially explicit rarefaction curves using different indices of taxonomic, functional and phylogenetic diversity. The user can also rarefy any biodiversity metric as provided by a self-written function (or an already existent one) that gives as output a vector with the values of a certain index of biodiversity calculated per plot (Ricotta, C., Acosta, A., Bacaro, G., Carboni, M., Chiarucci, A., Rocchini, D., Pavoine, S. (2019) <doi:10.1016/j.ecolind.2019.105606>; Bacaro, G., Altobelli, A., Cameletti, M., Ciccarelli, D., Martellos, S., Palmer, M. W., … Chiarucci, A. (2016) <doi:10.1016/j.ecolind.2016.04.026>; Bacaro, G., Rocchini, D., Ghisla, A., Marcantonio, M., Neteler, M., & Chiarucci, A. (2012) <doi:10.1016/j.ecocom.2012.05.007>).

Depends R (>= 3.5.0)

Imports ade4,adiv,ape,dplyr,geiger,methods,stats,vegan

Suggests kableExtra,knitr,phyloregion,picante,phytools,raster,rmarkdown

License GPL (>= 2)

Encoding UTF-8

VignetteBuilder knitr

NeedsCompilation no

Author Elisa Thouverai [aut, cre],
Sandrine Pavoine [aut],
Enrico Tordoni [aut],
Duccio Rocchini [aut],
Carlo Ricotta [aut],
Alessandro Chiarucci [aut],
Giovanni Bacaro [aut]

Maintainer Elisa Thouverai <`elisa.th95@gmail.com`>

Repository CRAN

Date/Publication 2021-03-11 23:10:09 UTC

R topics documented:

`directionalSAC` *Spatial Explicit Rarefaction Curves*

Description

The function calculates directional and non-directional accumulation curves of species diversity as a function of sampling effort.

Usage

```
directionalSAC(community,gradient)
```

Arguments

`community`	a community dataframe with N plots as rows, S species as columns. Both presence/absence and species abundances are allowed as entries. Plot names should be provided as row names on the dataframe.
`gradient`	a vector of numeric values, a matrix, a dataframe or an object of class `dist`.

Details

If `gradient` is a vector, then plots are ordered along a single spatial or environmental gradient. The length of the vector must be the number of rows in object community. Values in the vector must be in the same order as plots in object community. If names are given to each value of the vector, then the row names in `community` must be equal as those in `gradient`. For example, values in the vector may be the latitude of each plot or an environmental variable such as the temperature.

If `gradient` is a matrix or a data frame, then each column of the matrix must be a gradient along which plots are ordered. Only numeric values are allowed. For example, values in the matrix may be environmental data such as the temperature, the precipitation, the elevation. The result given by the function is an average over all gradients specified by the matrix. The number of rows in the matrix must be the same as the number of rows in `community`. If the matrix has row names, then they should be the same as in `community`. There must be as many columns in `gradient` as there are gradients of interest (for example, as many columns as there are environmental variables).

If `gradient` is of class `dist`, then the result given by the function is an average over all possible directional accumulation curves. In that case, `gradient` contains any pairwise dissimilarity/distance measure among plots. If names are given in `gradient`, they should be the same as row names in `community`. As specified in the main text, adjacent plots are combined step by step using the specified distance among plots as a constraining factor. In the simplest case, given a set of N plots, for each plot, the first, second, ..., k-th nearest neighbor are determined and a directional species accumulation curve is constructed using the resulting sequence of plots. This procedure is repeated for all plots, generating N directional accumulation curves from which a mean spatially explicit curve is calculated. The resulting curve is thus an intermediate solution between a non-directional accumulation curve and a pure directional curve in which all plots are ordered along a single spatial or environmental gradient.

Value

An object of class `data.frame` is returned containing the following statistics:

- N_SCR: Directional species accumulation curve.

- N_Exact: Non directional species accumulation curve (classic accumulation curve).

- Alpha_dir: directional mean number of species in the M plots (for details on the calculation, see Ricotta et al., 2019).

- Beta_M_dir: Directional beta diversity as a function of sampling effort M (for details on the calculation, Ricotta et al., 2019).

- Beta_N_dir: Normalized directional beta diversity.

- Beta_M: Non-directional beta diversity as a function of sampling effort M.

- Beta_N: Normalized non-directional beta diversity as a function of sampling effort M.

- Beta_Autocor: A normalized measure of autocorrelation for directional beta diversity calculated as the normalized difference between directional and non-directional beta.

Author(s)

Sandrine Pavoine `<sandrine.pavoine@mnhn.fr>`

Giovanni Bacaro `<gbacaro@units.it>`

References

Ricotta, C., Acosta, A., Bacaro, G., Carboni, M., Chiarucci, A., Rocchini, D., Pavoine, S. (2019) Rarefaction of beta diversity. *Ecological Indicators*, **107**, 105606. doi: 10.1016/j.ecolind.2019.105606.

Chiarucci, A., Bacaro, G., Rocchini, D., Ricotta, C., Palmer, M. W., Scheiner, S. M. (2009) Spatially constrained rarefaction: incorporating the autocorrelated structure of biological communities into sample-based rarefaction. *Community Ecology*, **10(2)**, 209–214.

Examples

```
require(vegan)
data(mite)
data(mite.xy)
comm_matrix <- mite
```

```
# Spatially-explicit curves can be obtained as follows

spatialdist <- dist(mite.xy) # to calculate the geographic
# distance between plots, i.e. the Euclidean distance
# between the coordinates of the plots)

betas <- directionalSAC(comm_matrix, spatialdist) # to calculate directional
# and non directional beta diversity

plot(1:70, betas$N_Exact, xlab="M", ylab="Species richness", ylim=range(c(betas$N_Exact,
betas$N_SCR, betas$Alpha, mean(apply(comm_matrix, 1, function(x) length(x[x>0]))))))
points(1:70,rep( mean(apply(comm_matrix, 1, function(x) length(x[x>0]))), 70), pch=2)
points(1:70, betas$N_SCR, pch=3)
points(1:70, betas$Alpha_dir, pch=4)
legend("right", legend=c("Non-directional SAC",
    "Non-directional alpha diversity", "Directional SAC",
    "Directional alpha diversity"), pch=1:4)
# M is the number of plots

plot(1:70, betas$Beta_M, xlab="M", ylab="Beta diversity",
 ylim=range(c(betas$Beta_M_dir, betas$Beta_M)))
points(1:70, betas$Beta_M_dir, pch=2)
legend("right", legend=c("Non-directional beta", "Directional beta"), pch=1:2)

plot(2:70, betas$Beta_N[2:70], xlab="M", ylab="Normalized beta diversity",
    ylim=range(c(betas$Beta_N_dir[2:70], betas$Beta_N[2:70])))
points(2:70, betas$Beta_N_dir[2:70], pch=2)
legend("right", legend=c("Non-directional beta", "Directional beta"), pch=1:2)

plot(2:70, betas$Beta_Autocor[2:70], xlab="M",
    ylab="Normalized measure of autocorrelation")
```

`duneFVG` *Abundance of Vascular Plants Data in Northen Adriatic Dunes*

Description

This data set is a list which contains information about the abundance of native and alien vascular plants sampled in 128 16 m^2 squared plots in the Northern Adriatic Dunes. See Tordoni et al. (2019) for details.

Usage

```
data("duneFVG")
```

Format

A list that contains 3 objects:

- `total` a numeric dataframe with plots as rows and species as columns. It contains the abundances of all the species collected.

`native` a numeric dataframe with plots as rows and species as columns. It contains the abundances of the native species collected.

`alien` a numeric dataframe with plots as rows and species as columns. It contains the abundances of the alien species collected.

There are two linked lists to `duneFVG`, one with the geographic coordinates of the plots (`duneFVG.xy`) and one with the functional traits of the species (`duneFVG.tr8`).

Source

Tordoni, E., Petruzzellis, F., Nardini, A., Savi, T., Bacaro, G. (2019) Make it simpler: alien species decrease functional diversity of coastal plant communities. *Journal of Vegetation Science*, **30**, 498–509.

Examples

```
data(duneFVG)
```

`duneFVG.tr8` *Functional Traits of Vascular Plants Data in Northern Adriatic Dunes*

Description

This dataset is a list which contains the functional traits of 71 vascular plant species sampled in 128 16 m^2 squared plots in the Northernn Adriatic Dunes. The species are the ones contained in the dataset `duneFVG`. See Tordoni et al. (2019) for details.

Usage

```
data("duneFVG.tr8")
```

Format

A list which contains 3 objects:

`traits.tot` a dataframe which contains 15 functional traits in columns and 71 species as rows. Species correspond to those contained in `duneFVG$total`.

`traits.nat` a dataframe which contains 15 functional traits in columns and 62 species as rows. Species correspond to those contained in `duneFVG$native`.

`traits.ali` a dataframe which contains 15 functional traits in columns and 9 species as rows. Species correspond to those contained in `duneFVG$alien`.

There are two linked lists to `duneFVG.tr8`, one with the abundance data of the species (`duneFVG`) and one with the spatial coordinates of the plots (`duneFVG.xy`).

Details

The following leaf functional and mechanistic traits were measured: specific leaf area (SLA, mm2 mg-1), major and minor vein length per unit area (VLAmaj VLAmin, mm mm^-2), osmotic potential at full turgor (π_0, -MPa), water potential at turgor loss point (Ψtlp, -MPa), N and C content (N % and C %), C and N stable isotope composition (δ13C per thousand, and δ15N per thousand) and C to N ratio (C:N).

The following categorical traits associated with phenology and reproductive strategies derived from literature are also included: life form, clonality, photosynthetic pathway, sprouting and flowering time-span.

Source

Tordoni, E., Petruzzellis, F., Nardini, A., Savi, T., Bacaro, G. (2019) Make it simpler: alien species decrease functional diversity of coastal plant communities. *Journal of Vegetation Science*, **30**, 498–509.

Examples

```
data(duneFVG.tr8)
```

`duneFVG.xy` *Geographical Coordinates of Plots, in Northen Adriatic Dunes*

Description

This data set is a list that contains the geographic coordinates of 128 16 m^2 squared plots in the Northern Adriatic Dunes where the alien and native vascular plants of the data set `duneFVG` were collected. See Tordoni et al. (2019) for details.

Usage

```
data("duneFVG.xy")
```

Format

A list that contains 3 objects:

- `tot.xy` a numeric dataframe that contains the coordinates of 128 plots in the Northern Adriatic Dunes where all the species of `duneFVG$total` are collected.
- `nat.xy` a numeric dataframe that contains the coordinates of 127 plots in the Northern Adriatic Dunes where the native species of `duneFVG$native` are collected.
- `ali.xy` a numeric dataframe that contains the coordinates of 106 plots in the Northern Adriatic Dunes where the alien species of `duneFVG$alien` are collected.

There are two linked lists to `duneFVG.xy`, one with the abundance data of the species (`duneFVG`) and one with the functional traits of the species (`duneFVG.tr8`).

Source

Tordoni, E., Petruzzellis, F., Nardini, A., Savi, T., Bacaro, G. (2019) Make it simpler: alien species decrease functional diversity of coastal plant communities. *Journal of Vegetation Science*, **30**, 498–509.

Examples

```
data(duneFVG.xy)
```

`rao_permuted` *Functional Rarefaction Curves on Standardized Species Number*

Description

The function calculates functional rarefaction curves using Rao's quadratic entropy (Q) for a reference community. However, functional distance matrixes are calculated by selecting randomly species with their related functional traits from another larger community. The number of species randomly selected corresponds to the number of species in the reference community matrix. The relative abundances of the standard community are used instead for the iterative calculation of the rarefacted index.

Usage

```
rao_permuted(comm_st,dist_f,random=99)
```

Arguments

`comm_st`	a dataframe or matrix of the standard community with N plots as rows, S species as columns. Only the abundances of species in plots are allowed as entries. Plot names should be provided as row names.
`dist_f`	an object of class `dist` containing the pairwise functional distances among the species of the larger community. The number of species must be larger than `comm_st`. The distance matrix should be squared Euclidean or simply Euclidean.
`random`	number of iterations for the selection of species from the whole community. The functional rarefaction curve and associated Confidence Intervals are are calculated as the average of these iterations.

Details

Rao's quadratic entropy (Rao 1982) is commonly used as a measure of functional diversity. It incorporates both the relative abundance of species with their pairwise functional dissimilarities. It expresses the average difference between two randomly selected individuals with replacements.

In the case study below, the functional dissimilarity matrix is calculated using the function `dist.ktab` from the package `ade4`. A generalization of the Gower's general coefficient of distance is used to allow the treatment of various statistical types of variables when calculating the functional pairwise distances of the species.

To calculate the rarefaction curve, a standard number of species that corresponds to the number of species of `comm_st` are randomly selected in `tr8` with their related functional traits, then a functional distance matrix is obtained and Q is calculated for 1,2,...,N plots using the mean relative abundances of the species in `comm_st`. This procedure is repeated `random` times and the average rarefaction curve is calculated.

Value

An object of class `data.frame` with 3 columns is returned:

- Rarefaction: : mean of the values of the accumulation curves for all the sampling dimensions;

- IC_up: upper confidence interval;

- IC_low: lower confidence interval.

Author(s)

Elisa Thouverai `<elisa.th95@gmail.com>`

Enrico Tordoni `<etordoni@units.it>`

References

Botta-Dukat, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, **16**, 533–540.

Pavoine S., Vallet, J., Dufour, A.-B., Gachet, S. and Daniel, H. (2009) On the challenge of treating various types of variables: Application for improving the measurement of functional diversity. *Oikos*, **118**, 391–402.

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.

Ricotta, C., Pavoine, S., Bacaro, G., Acosta, A. (2012) Functional rarefaction for species abundance data. *Methods in Ecology and Evolution*, **3**, 519–525.

Tordoni, E., Petruzzellis, F., Nardini, A., Savi, T., Bacaro, G. (2019) Make it simpler: alien species decrease functional diversity of coastal plant communities. *Journal of Vegetation Science*, **30**, 498—509.

See Also

`ser_functional`

Examples

```
require(ade4)
data(duneFVG)
data(duneFVG.tr8)
tr8_N<-duneFVG.tr8$traits.tot[,c(1,3,4)]
tr8_D<-data.frame(duneFVG.tr8$traits.tot[,2])
tr8_Q<-duneFVG.tr8$traits.tot[,5:15]
tr8dist<-dist.ktab(ktab.list.df(list(tr8_N,tr8_D,tr8_Q)),type=c('N','D','Q'))
tr8dist<-cailliez(tr8dist)
## Rarefaction
```

```
rare<-rao_permuted(duneFVG$alien,tr8dist)
plot(rare$Rao, ylab="Rao QE", xlab="Number of plots", type="l", ylim=range(rare))
lines(rare[,2], lty=2)
lines(rare[,3], lty=2)
```

`rarexpected_fun` *Functional Diversity Rarefaction Curves*

Description

The function calculates functional diversity rarefaction curves using the method proposed by Ricotta et al. (2010).

Usage

```
rarexpected_fun(comm,dist_f)
```

Arguments

`comm`	a community data frame or matrix with N plots as rows, S species as columns. Only presence/absence (1/0) of species in plots are allowed as entries.
`dist_f`	an object of class `dist` containing the pairwise functional distances among species. The distance matrix should be squared Euclidean or simply Euclidean.

Details

The expected occupancy frequencies for species at an increasingly larger cumulative number of plots N are used for the calculation of Rao's quadratic entropy to obtain an expected functional diversity between species in the N random plots. The expected occupancy frequency is the probability for every species in the community to find that species in the N random plots.

Value

A vector containing the values of the expected functional diversity for every sampling dimension is returned.

Author(s)

Giovanni Bacaro <`gbacaro@units.it`>

Elisa Thouverai <`elisa.th95@gmail.com`>

References

Botta-Dukat, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, **16**, 533–540.

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.

Ricotta, C., Burrascano, S., Blasi, C. (2010) Incorporating functional dissimilarities into sample-based rarefaction curves: from taxon resampling to functional resampling. *Journal of Vegetation Science*, **21**, 280–286.

See Also

`ser_functional`

Examples

```
require(ade4)
require(vegan)
data(duneFVG)
data(duneFVG.tr8)
tr8_N<-duneFVG.tr8$traits.tot[,c(1,3,4)]
tr8_D<-data.frame(duneFVG.tr8$traits.tot[,2])
tr8_Q<-duneFVG.tr8$traits.tot[,5:15]
tr8dist<-dist.ktab(ktab.list.df(list(tr8_N,tr8_D,tr8_Q)),type=c('N','D','Q'))
tr8dist<-cailliez(tr8dist)
d<-decostand(duneFVG$total, method='pa')
rare<-rarexpected_fun(d,tr8dist)
plot(rare, ylab="Rao QE", xlab="Number of plots", type="l")
```

`rare_alpha` *Alpha Diversity Rarefaction Curve*

Description

The function calculates spatial and non-spatially explicit rarefaction curves using different indexes of species alpha diversity.

Usage

```
rare_alpha(comm,dist_xy=NULL,method=c("HCDT","hill","fun_div"),q=0,random=99,fun_div=NULL
,args=NULL,verbose=FALSE,spatial=FALSE,mean=FALSE)
```

Arguments

`comm`	a community data frame or matrix with N plots as rows, S species as columns. Both the presence/absence (1/0) or the abundances of species in plots are allowed as entries. Plot names should be provided as row names.

`dist_xy`	an object of class `dist` containing the pairwise geographic distances among the plots. The names of the labels must be the same as the rows of `comm`. The argument could be empty (default=NULL) if non-spatial rarefaction curves are considered.
`method`	the diversity index for the calculation of the rarefaction curve, one among "HCDT","hill" or "fun_div". See details.
`q`	parameter that determines the sensitivity of the measure to the relative abundance of the species (ignored if method = fun_div).
`random`	number of iterations to construct the non-spatially explicit rarefaction curve
`fun_div`	a string with the name of the user-defined function for the diversity index in the rarefaction. The function must calculate the value of the chosen diversity index per plot and return a numeric vector with the values calculated.
`args`	a list with the arguments for fun_div. The value NA should be given in place of the community matrix in the list. The names of the elements must correspond to the names of the arguments of the function passed.
`verbose`	if TRUE, the arguments of `fun_div` are inserted interactively by the user.
`spatial`	if TRUE, the function calculates the spatially-explicit rarefaction.
`mean`	if TRUE, the mean of the relative abundances of the species in all the sampling dimensions is used for the calculation of the rarefaction instead of the total abundance, giving less importance to abundant species.

Details

The available methods are:

`HCDT`: HCDT entropy is a generalization of the standard coefficient of entropy (Harvda and Charvat 1967; Daroczy 1970; Tsallis 1988):

$$HCDT = \frac{1 - \left(\sum_i^S (p_i)^q\right)}{q-1}$$

where S is the number of species, p_i is the relative abundance of species i and q is the parameter that regulates the sensitivity to species abundance. When q tends to 1, the equation becomes the Shannon entropy:

$$HCDT = -\sum_{i=1}^{S} p_i log(p_i)$$

`hill`: Hill numbers (${}^qD$) is a class of measures that obeys to the replication principle and integrates species richness and species abundances. The parameter q, called 'order', regulates the sensitivity of the index to the species abundance: with $q = 0$ the value of the index correspond to the species richness, with $q = 1$ the measure tends to the exponential of Shannon index, and with $q = 2$ it corresponds to the inverse of Simpson index:

$${}^q\mathrm{D} = \left(\sum_{i=1}^{S} p_i^q\right)^{1/(1-q)}$$

where S is the number of species, p_i is the relative abundance of species i and q is the order.

`hill` doesn't behave like a non-decreasing function of the increasing number of plots *N*. This behavior is typical of non-concave indexes.

Value

An object of class `data.frame` with 3 columns is returned:

- Rarefaction: mean of the values of the accumulation curves for all the sampling dimensions;

- IC_up: upper confidence interval;

- IC_low: lower confidence interval.

Author(s)

Elisa Thouverai <`elisa.th95@gmail.com`>

with contributions of Sandrine Pavoine.

References

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., Ellison, A. M (2014). Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, **84**, 45–67.

Daroczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.

Havrda, M., Charvat F. (1967) Quantification method of classification processes: concept of structural alpha-entropy. *Kybernetik*, **3**, 30–35.

Ricotta, C., Pavoine, S., Bacaro, G., Acosta, A. (2012) Functional rarefaction for species abundance data. *Methods in Ecology and Evolution*, **3**, 519–525.

Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 480–487.

Examples

```
## Not run:
#Time consuming
require(adiv)
data(duneFVG)
data(duneFVG.xy)
dist_sp<-dist(duneFVG.xy$tot.xy,'euclidean')
## Spatial explicit rarefaction with HCDT
rare<-rare_alpha(duneFVG$total,dist_sp,q=2,random=99,spatial=TRUE)
## Non-spatially explicit rarefaction with Hill
rare1<-rare_alpha(duneFVG$total,dist_sp,method='hill',q=2,mean=TRUE)
##fun_div with the list of arguments passed
#to set the argument, digit the correspondent number
#digit the number correspondent to "stop" when the process is complete
#remember to load the package that contain the function you want to exploit
#Example with the function speciesdiv() of adiv package
a<-list(NA,'Shannon')
names(a)<-c('comm','method')
rare2<-rare_alpha(duneFVG$total,method="fun_div",random=999,fun_div='speciesdiv',args=a,mean
=TRUE)
rare3<-rare_alpha(duneFVG$total,dist_sp,method="fun_div",fun_div='speciesdiv',args=a,spatial
=TRUE,mean=TRUE)
##fun_div without the list of arguments passed (verbose=TRUE)
```

```
rare4<-rare_alpha(duneFVG$total,method="fun_div",random=999,fun_div='speciesdiv',verbose=TRUE,mean
=TRUE)
rare5<-rare_alpha(duneFVG$total,dist_sp,method="fun_div",fun_div='speciesdiv',verbose=TRUE,spatial
=TRUE,mean=TRUE)

## End(Not run)
```

`rare_beta` *Beta Diversity Rarefaction Curve*

Description

The function calculates spatial and non-spatially explicit rarefaction curves using different indexes of species beta diversity.

Usage

```
rare_beta(comm,dist_xy=NULL,method=c("whittaker","jaccard","bray","cody"
,"fun_div"),random=99,fun_div=NULL,args=NULL,verbose=FALSE,spatial=FALSE)
```

Arguments

`comm`	a community dataframe or matrix with N plots as rows, S species as columns. Both the presence/absence (1/0) or the abundances of species in plots are allowed as entries. Plot names should be provided as row names.
`dist_xy`	an object of class `dist` containing the pairwise geographic distances among the plots. The names of the labels must be the same as the rows of `comm`. The argument could be empty (default=NULL) if non-spatial rarefaction curves are considered.
`method`	the diversity index for the calculation of the rarefaction curve, one among "whittaker","jaccard","bray","cody" (only for spatially-explicit rarefaction) or "fun_div". See details.
`random`	number of iterations to construct the non-spatially explicit rarefaction curve
`fun_div`	a string with the name of the function you want to use to calculate the rarefaction. The function must calculate the value of the chosen index per plot and return a pairwise distance matrix with the values.
`args`	a list with the arguments for fun_div. The value `NA` should be given in place of the community matrix in the list. The names of the elements must correspond to the names of the arguments of the function passed.
`verbose`	if TRUE, the arguments of `fun_div` are inserted interactively by the user.
`spatial`	if TRUE, the function calculates the spatially-explicit rarefaction.

Details

The available methods are:

`whittaker`: Whittaker's species turnover calculates how many times there is a change in species composition among the plots:

$$\beta_w = \gamma/\alpha - 1$$

where γ is the species richness over all plots compared and α the average species richness within a single plot.

`jaccard`: Jaccard dissimilarity coefficient is a pairwise measure of dissimilarity between plots:

$$\beta_j = a/(\alpha_1 + \alpha_2 - a)$$

where a is the number of species in common between two plots, and α_1 and α_2 are the values of alpha diversity (species richness) of the 2 plots compared. The rarefaction curve is calculated with the mean of the pairwise diversities in N random plots.

`bray`: Bray-Curtis dissimilarity is a pairwise measure of similarity between plots weighted by the abundances of the species:

$$\beta_{bray} = \frac{\sum_i |x_i - x_j|}{\sum_i x_i + x_j}$$

where x_i and x_j are the abundances of the species x in the plots i and j.

`cody`: Cody index is defined as the rate at which species are being replaced in censuses at each point on the habitat gradient and is fixed for samples arranged along gradients of environmental change:

$$\beta_c = [g(H) + l(H)]/2$$

where $g(H)$ is the number of species gained along the habitat gradient H and $l(H)$ is the number of species lost.

In some cases, the rarefaction doesn't behave like a non-decreasing function of plot number *N*. This behavior is typical of non-concave indexes.

Value

An object of class `data.frame` with 3 columns is returned:

- Rarefaction: : mean of the values of the accumulation curves for all the sampling dimensions;

- IC_up: upper confidence interval;

- IC_low: lower confidence interval.

Author(s)

Elisa Thouverai <`elisa.th95@gmail.com`>

Enrico Tordoni <`etordoni@units.it`>

with contributions of Sandrine Pavoine.

References

Baselga, A. (2013) Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. *Ecography*, **36**, 124–128.

Koleff, P., Gaston, K.J., Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.

Wilson, M.V., Shmida, A. (1984) Measuring Beta Diversity with Presence-Absence Data. *Journal of Ecology*, **72**, 1055–1064.

Examples

```
## Not run:
#Time consuming
data(duneFVG)
data(duneFVG.xy)
dist_sp<-dist(duneFVG.xy$tot.xy,'euclidean')
## Spatial explicit rarefaction
## Take care that the following scripts may be a bit time-consuming
## depending on your computer (less than a minute to a few minutes)
rare<-rare_beta(duneFVG$total,dist_sp,method="bray",spatial=TRUE)
rare1<-rare_beta(duneFVG$total,method="bray",random=99)
plot(rare [,1], ylab="Rao QE", xlab="Number of plots", type="l", ylim=range(rare, na.rm
=TRUE))
lines(rare[,2], lty=2)
lines(rare[,3], lty=2)
plot(rare1 [,1], ylab="Rao QE", xlab="Number of plots", type="l", ylim=range(rare, na.rm
=TRUE))
lines(rare1[,2], lty=2)
lines(rare1[,3], lty=2)


## End(Not run)
```

`rare_phylo` *Phylogenetic Diversity Rarefaction Curves*

Description

`rare_phylo` calculates classic rarefaction curves using different indexes of phylogenetic diversity.

`ser_phylo` calculates spatially-explicit rarefaction curves using different indexes of phylogenetic diversity.

Usage

```
rare_phylo(comm,tree=NULL,method=c("faith","barker","Ia","hill","tsallis",
"renyi","fun_div"),exp=0,resampling=99,fun_div=NULL,args=NULL,verbose=FALSE)

ser_phylo(comm,tree=NULL,dist_xy,method=c("faith","barker","Ia","hill","tsallis"
,"renyi","fun_div"),exp=0,fun_div=NULL,args=NULL,verbose=FALSE,
comparison=FALSE,resampling=99)
```

Arguments

`comm`	a community dataframe or matrix with N plots as rows, S species as columns. Both the presence/absence (1/0) or the abundances of species in plots are allowed as entries. Plot and species names should be provided as row names and column names.
`tree`	an object of class `phylo` or `phylo4` containing the phylogenetic tree of the species which are present in `comm`.
`dist_xy`	an object of class `dist` containing the pairwise geographic distances among the plots or an object of class, an object of class `vector` containing the order of the sampling units along a gradient or an object of class `matrix` or `data.frame` where each column represents a gradient along which the sampling units are ordered. The names of the labels must be the same as the rows of `comm`. If `dist_xy` is not of class dist, then a distance matrix is calculated using the Euclidean distance coefficient.
`method`	the diversity index for the calculation of the rarefaction curve, one among "faith", "barker", "Ia", "hill", "tsallis", "renyi" or "fun_div". See details.
`exp`	parameter that determines the sensitivity of the measure to the relative abundance of the species for "Ia", "hill", "tsallis" and "renyi" indexes.
`resampling`	number of times plots (rows) are randomly resampled from comm to calculate the mean accumulation curve for the non-spatially-explicit rarefaction.
`fun_div`	a string with the name of the user-defined function for the diversity index in the rarefaction. The function must calculate the value of the chosen diversity index per plot and return a numeric vector with the values calculated.
`args`	a list with the arguments for fun_div. The value NA should be given in place of the community matrix in the list. The names of the elements must correspond to the names of the arguments of the function passed.
`verbose`	if TRUE, the arguments of `fun_div` are inserted interactively by the user.
`comparison`	if TRUE, both non-spatially explicit and spatially explicit phylogenetic rarefactions are calculated.

Details

The available methods are:

`faith`: Faith's phylogenetic diversity (PD) is defined as the sum of branch lengths in a phylogenetic tree for the assemblage of species (Faith 1992):

$$PD = \sum_{i \in B} L_i$$

where L_i is the branch length of the branch i and B is the number of branches in the tree.

`barker`: Barker's index is the abundance weighted Faith's PD. The number of branches (B) is multiplied by the weighted mean branch length, with weights equal to the average abundance of species sharing that branch (Vellend et al. 2010):

$$PDw = B \times \frac{\sum_i^B L_i A_i}{\sum_i^B A_i}$$

where L_i is the branch length of the branch i, and A_i is the average abundance of the species sharing the branch i. B is the number of branches in the tree.

`Ia`: Ia index, by Pavoine et al. (2009), calculates PD partitioned between evolutionary periods and between plots defined in terms of spatial and time units. Tsallis or HCDT entropy (Harvda and Charvat 1967; Daroczy 1970; Tsallis 1988) (that measures diversity by regrouping individuals into categories) is computed for each period of the phylogenetic tree, from the number of lineages that descend from the period and from the relative abundances summed within these lineages within the focal community. With $exp = 0$, HCDT is the richness (number of species) minus one and Ia is Faith's PD minus the height of the phylogenetic tree; with exp tending to 1 HCDT is a generalization of the Shannon index while with $exp = 2$ HCDT is the Simpson index and Ia is Rao's QE applied to phylogenetic distances between species. To apply Ia, the phylogeny must be ultrametric:

$$Ha = \frac{(1 - \sum_{i=1}^{n} p_i^a)}{(a-1)}$$

the equation for the HCDT entropy, where p_i is the relative abundance of the species i and a is the scaling constant that weights the importance of rarity of the species.

$$I_a = \sum_{K=1}^{N} (t_K - t_{K-1}) H_{a,K}$$

where $H_{a,K}$ is Ha applied to the period K and $t_K - t_{K-1}$ is the length of the period K.

Hill index (`hill`) and the HCDT (`tsallis`) and Renyi (`renyi`) entropies are adapted for the calculation of the phylogenetic diversity replacing the species with the units of the branch length in the phylogenetic tree (Pavoine & Ricotta 2019):

$$Hill = \left[\sum_{i \in B} L_i (p_i)^q] \right]^{\frac{1}{1-q}}$$

$$HCDT = \frac{1 - \sum_{i \in B} L_i (p_i)^q}{q - 1}$$

$$Renyi = \frac{1}{1-q} log \left[\sum_{i \in B} L_i (p_i)^q \right]$$

where L_i is the branch length of the branch i, p_i is the relative abundance of the species sharing the branch i and q is the scaling constant that weights the importance of rarity of the species. B is the number of branches in the tree.

Value

An object of class `data.frame` with 3 columns is returned:

- Rarefaction: : mean of the values of the accumulation curves for all the sampling dimensions;

- IC_up: upper confidence interval;

- IC_low: lower confidence interval.

If comparison is TRUE, the `data.frame` object will have six columns, with the values of the accumulation curve and confidence intervals for both spatially explicit and non-spatially explicit rarefaction.

Author(s)

Elisa Thouverai <`elisa.th95@gmail.com`>

with contributions of Sandrine Pavoine.

References

Chao, A., Chiu, C.-H., Hsieh, T.C., Davis, T., Nipperess, D.A., Faith, D.P. (2014) Rarefaction and extrapolation of phylogenetic diversity. *Methods in Ecology and Evolution*, **6**, 380–388.

Daroczy, Z. (1970) Generalized information functions. *Information and Control*, **16**, 36–51.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.

Havrda, M., Charvat F. (1967) Quantification method of classification processes: concept of structural alpha-entropy. *Kybernetik*, **3**, 30–35

Pavoine, S., Love, M., Bonsall, M.B. (2009) Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: Sebastes) in the Southern California Bight. *Ecology Letters*, **12**, 898–908.

Pavoine, S., Ricotta, C. (2019) A simple translation from indices of species diversity to indices of phylogenetic diversity. *Ecological Indicators*, **101**, 552–561.

Swenson, N.G. (2014) Functional and Phylogenetic Ecology in R. Springer UseR! Series, Springer, New York, New York, U.S.A.

Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 480–487.

Vellend, M., Cornwell, W.K., Magnuson-Ford, K., Mooers, A. (2010) Measuring phylogenetic biodiversity. *Magurran & McGill*, 194–207.

Examples

```
## Not run:
#Time consuming
require(picante)
require(geiger)
data(phylocom)
phylo<-treedata(phylocom$phylo,phylocom$sample[1,],warnings = FALSE)$phy
## Non-spatially explicit rarefaction
raref<-rare_phylo(phylocom$sample,phylo,resampling=999) ##Faith index
plot(raref [,1], ylab="Faith", xlab="Number of plots", type="l", ylim=range(raref, na.rm
=TRUE))
lines(raref[,2], lty=2)
lines(raref[,3], lty=2)
rareb<-rare_phylo(phylocom$sample,phylo,method='barker',resampling=999) ##Barker index
plot(rareb [,1], ylab="Barker", xlab="Number of plots", type="l", ylim=range(rareb, na.rm
=TRUE))
lines(rareb[,2], lty=2)
lines(rareb[,3], lty=2)
rareia<-rare_phylo(phylocom$sample,phylo,method='Ia',resampling=999,exp=2) ##Ia index
plot(rareia [,1], ylab="Ia", xlab="Number of plots", type="l", ylim=range(rareia, na.rm
=TRUE))
```

```
lines(rareia[,2], lty=2)
lines(rareia[,3], lty=2)

## End(Not run)
```

`ser_functional` *Functional Spatially-Explicit Rarefaction Curves*

Description

The function calculates classic and spatially-explicit rarefaction curves using different indexes of functional diversity.

Usage

```
ser_functional(comm,dist_f=NULL,dist_xy,method=c("rao","chao","fun_div"),
tau=NA,q=0,comparison=FALSE,resampling=99,fun_div=NULL,args=NULL,verbose=FALSE)
```

Arguments

`comm`	a community dataframe or matrix with N plots as rows, S species as columns. Only the abundances of species in plots are allowed as entries for the methods `rao` and `chao`. Plot names should be provided as row names.
`dist_f`	an object of class `dist` containing the pairwise functional distances among species. The distance matrix should be squared Euclidean or simply Euclidean.
`dist_xy`	an object of class `dist` containing the pairwise geographic distances among the plots or an object of class, an object of class `vector` containing the order of the sampling units along a gradient or an object of class `matrix` or `data.frame` where each column represents a gradient along which the sampling units are ordered. The names of the labels must be the same as the rows of `comm`. If `dist_xy` is not of class dist, then a distance matrix is calculated using the Euclidean distance coefficient.
`method`	the diversity index for the calculation of the rarefaction curve, one among "rao", "chao" or "fun_div". See details.
`tau`	threshold of functional distinctiveness between any two species for `chao`. See details.
`q`	parameter that determines the sensitivity of the measure to the relative abundance of the species for `chao`.
`comparison`	if TRUE, both non-spatially explicit and spatially explicit rarefactions are calculated.
`resampling`	number of times plots (rows) are randomly resampled from comm to calculate the mean accumulation curve for the non-spatially-explicit rarefaction.
`fun_div`	a string with the name of the user-defined function for the diversity index in the rarefaction. The function must calculate the value of the chosen diversity index per plot and return a numeric vector with the values calculated.

`args`	a list with the arguments for fun_div. The value NA should be given in place of the community matrix in the list. The names of the elements must correspond to the names of the arguments of the function passed.
`verbose`	if TRUE, the arguments of `fun_div` are inserted interactively by the user.

Details

The available methods are:

`rao`: Rao functional diversity is based on the quadratic entropy of Rao (1982) and incorporates both the relative abundances of species and a measure of the pairwise functional distances between species. It expresses the average difference between two randomly selected individuals with replacements.

$$Q(\mathbf{p}_i, \mathbf{D}) = \sum_{k=1}^{S}\sum_{l=1}^{S} p_k p_l d_{kl}$$

where $p = (p1, ..., p_k, ..., S)$ is the vector of relative species abundances, S is the number of species, $\mathbf{D} = (d_{kl})$ is the matrix of functional dissimilarities among species, and d_{kl} is the functional dissimilarity between species k and l.

`chao`: Chao et al. (2018) functional beta-diversity index FD quantifies the effective number of equally-distinct functional groups in the considered plots at the distinctiveness `tau` threshold. Any two species with functional distance greater than or equal to `tau`, are treated as functionally equally-distinct and as belonging to different functional groups with distance `tau`. For each pair of species with functional distance lower than `tau` but different from zero, only a proportion of individuals is considered functionally equally-distinct, the other proportion of individuals is considered functionally indistinct. If the pairwise distance is equal to zero, the two species are treated as belonging to the same functional group. After dividing the set of species to form functionally indistinct groups, the contribution of every species is quantified and then the FD of order q is calculated using the Hill number of order q:

$${}^{q}\mathrm{FD}(\Delta(\tau)) = \left(\sum_{i=1}^{S} \nu_i(\tau)\left(\frac{a_i(\tau)}{n_+}\right)^{(1/1-q)}\right)$$

where $a_i(\tau)$ is the combined abundance of all functionally-indistinct individuals from species i, $v_i(\tau) = n_i/a_i(\tau)$ represents the attribute contribution of species i for a threshold level `tau` (n_i is the abundance of species i), n_+ is the total number of individuals in the community and q is the parameter that determines the sensitivity of the measure to the relative abundance of the species.

Value

An object of class `data.frame` with 3 columns is returned:

- Rarefaction: : mean of the values of the accumulation curves for all the sampling dimensions;

- IC_up: upper confidence interval;

- IC_low: lower confidence interval.

Author(s)

Elisa Thouverai `<elisa.th95@gmail.com>`

with contributions of Sandrine Pavoine.

References

Botta-Dukat, Z. (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, **16**, 533–540.

Chao, A., Chiu, C.-H., Villeger, S., Sun, I.F., Thorn, S., Lin, Y., Chiang, J.-M., B. Sherwin, W. (2018) An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs*, **89**, e01343.

Rao, C.R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.

Ricotta, C., Pavoine, S., Bacaro, G., Acosta, A. (2012) Functional rarefaction for species abundance data. *Methods in Ecology and Evolution*, **3**, 519–525.

See Also

`rare_Rao`

`rare_alpha`

Examples

```
## Not run:
#Time consuming
require(ade4)
data(duneFVG)
data(duneFVG.xy)
data(duneFVG.tr8)
tr8_N<-duneFVG.tr8$traits.tot[,c(1,3,4)]
tr8_D<-data.frame(duneFVG.tr8$traits.tot[,2])
tr8_Q<-duneFVG.tr8$traits.tot[,5:15]
tr8dist<-dist.ktab(ktab.list.df(list(tr8_N,tr8_D,tr8_Q)),type=c('N','D','Q'))
tr8dist<-cailliez(tr8dist)
dist_sp<-dist(duneFVG.xy$tot.xy,'euclidean')
## Rao index
## Take care that the following scripts may be time-consuming
rare<-ser_functional(duneFVG$total,tr8dist,dist_sp,method='rao',comparison=TRUE)
plot(rare [,1], ylab="Rao QE", xlab="Number of plots", type="l", ylim=range(rare, na.rm=TRUE))
lines(rare[,2], lty=2)
lines(rare[,3], lty=2)
lines(rare[,4 ], col=2)
lines(rare[,5], lty=2, col=2)
lines(rare[,6], lty=2, col=2)
legend("bottomright",  legend=c("spatially-explicit rarefaction", "classic rarefaction"),
lty=1, col=1:2)
# dashed lines represent the confidence interval
## Chao index
rareC<-ser_functional(duneFVG$total,tr8dist,dist_sp,method='chao',tau=0.6,q=2,comparison
=TRUE)
plot(rareC [,1], ylab="Chao", xlab="Number of plots", type="l", ylim=range(rareC, na.rm
=TRUE))
lines(rareC[,2], lty=2)
lines(rareC[,3], lty=2)
lines(rareC[,4 ], col=2)
```

```
lines(rareC[,5], lty=2, col=2)
lines(rareC[,6], lty=2, col=2)
legend("bottomright", legend=c("spatially-explicit rarefaction", "classic rarefaction"),
lty=1, col=1:2)
# dashed lines represent the confidence interval

## End(Not run)
```

Index