Package 'iNEXT'

January 28, 2020

Type Package

Title Interpolation and Extrapolation for Species Diversity

Version 2.0.20

Author T. C. Hsieh, K. H. Ma and Anne Chao

Description Provides simple functions to compute and plot two types (sample-size- and coverage-based) rarefaction and extrapolation of species diversity (Hill numbers) for individual-based (abundance) data or sampling-unitbased (incidence) data. (Hsieh, Ma and Chao 2014) <doi: 10.1111/2041-210X.12613>.

License GPL (>= 3)

URL http://chao.stat.nthu.edu.tw/wordpress/software_download/

Maintainer T. C. Hsieh <euler96@gmail.com>

Depends R (>= 3.4)

Imports stats, graphics, ggplot2, reshape2

BugReports https://github.com/JohnsonHsieh/iNEXT/issues

LazyLoad yes RoxygenNote 6.1.1

Suggests testthat, knitr, rmarkdown, gridExtra, ggthemes

VignetteBuilder knitr

ByteCompile true

NeedsCompilation no

Repository CRAN

Date/Publication 2020-01-28 13:10:04 UTC

R topics documented:

iNEXT-package															•		 	•							2
ant		•		•	•	•		•		•	•	•	•		•	 	 	•	•			•	•	•	3
as.abucount		•		•	•	•		•		•	•	•	•		•	 	 	•	•			•	•	•	4
as.incfreq								•		•				•	•		 	•							4

iNEXT-package

bi	rd	•	•	• •	 •	•	•	•			•	•	•	 •			•		•	•		•	•	•	•	•	•	•	5
Cł	naoRichness			•	 •								•																6
Cł	naoShannon		•	•	 •			•					•						•	•									7
Cł	naoSimpson		•	•	 •			•					•						•	•									8
cil	iates	•		•	 •						•	•	•	 •	•				•	•	•	•			•	•	•	•	9
Da	ataInfo		•	•	 •			•					•						•	•									10
es	timateD		•	•	 •			•					•						•	•									10
fo	rtify.iNEXT		•	•	 •			•					•						•	•									11
gg	iNEXT		•	•	 •			•					•						•	•									12
iN	ЕХТ		•	•	 •			•					•						•	•									13
ple	ot.iNEXT .		•	•	 •			•					•						•	•									15
pr	int.iNEXT .			•	 •								•	 •															16
sp	ider		•	•	 •						•	•	•		•				•	•	•	•			•	•	•	•	16
																													18

Index

iNEXT-package

Interpolation and extrapolation for species diversity

Description

iNEXT (iNterpolation and EXTrapolation) provides functions to compute and plot two types (samplesize- and coverage-based) interpolation and extrapolation sampling curves of Hill numbers for three most widely used members of Hill numbers (species richness, Shannon diversity and Simpson diversity) for individual-based abundance data or sampling-unit-based incidence data. iNEXT also computes bootstrap confidence intervals around the diversity for rarefied/extrapolated samples, facilitating the comparisons of diversities across multiple assemblages. The estimated asymptote along with a confidence interval for each of the three diversity measures is also provided. An auxiliary function is included to compute/compare diversities across multiple assemblages for a particular user-specified sample size or sample coverage. The sample-size-based rarefaction and extrapolation for species richness were developed by Colwell et al. (2012) and the corresponding coverage-based methodologies were developed by Chao and Jost (2012). Chao et al. (2014) extended the previous work for species richness to Hill numbers. The statistical methods and tools provided in iNEXT efficiently use all data to make more robust and detailed inferences about the sampled assemblages, and also to make objective comparisons of multiple assemblages. Online user guide/tutorials and iNEXT online version (Hsieh et al. 2013, 2015) are also available.

Author(s)

T. C. Hsieh K. H. Ma Anne Chao Maintainer: T. C. Hsieh <euler96@gmail.com>

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.

Chao, A. & Jost, L. (2012) Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology, 93, 2533-2547.

Colwell, R.K., Chao, A., Gotelli, N.J., Lin, S.-Y., Mao, C.X., Chazdon, R.L. & Longino, J.T. (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. Journal of Plant Ecology, 5, 3-21.

Hsieh, T.C., Ma, K.H. & Chao, A. (2013) iNEXT online: interpolation and extrapolation (Version 1.0) [Software]. Available from http://chao.stat.nthu.edu.tw/blog/software-download/.

Hsieh, T.C., Ma, K.H. & Chao, A. (2016) iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution (in revision).

ant

Ant species incidence frequencies data

Description

Ant species incidence frequencies for samples from five elevations in northeastern Costa Rica (Longino and Colwell 2011). The number of sampling units (1m x 1m forest floor plot) for the 5 sites are respectively 599, 230, 150, 200 and 200. The number of observed species for the 5 sites are respectively 227, 241, 122, 56 and 14.

Usage

data(ant)

Format

The input format for each site is a list of integers. For incidence data, the first entry must be the total number of sampling units, followed by the species incidence frequencies as shown below: List of 5

\$ h50m : num [1:228] 599 1 1 1 1 1 1 1 1 1 1 ... \$ h500m : num [1:242] 230 1 1 1 1 1 1 1 1 1 1 ... \$ h1070m: num [1:123] 150 1 1 1 1 1 1 1 1 1 ... \$ h1500m: num [1:57] 200 1 1 1 1 1 1 1 1 1 ... \$ h2000m: num [1:15] 200 1 2 2 3 4 8 8 13 15 ...

References

Longino, J.T. & Colwell, R.K. (2011) Density compensation, species composition, and richness of ants on a neotropical elevational gradient. Ecosphere, 2, art29.

as.abucount

Description

as.abucount: transform abundance raw data (a species by sites matrix) to abundance rwo-sum counts data (iNEXT input format).

Usage

as.abucount(x)

Arguments

х

a data.frame or matirx of species by sites matrix.

Value

a vector of species abundance row-sum counts.

Examples

```
data(ciliates)
lapply(ciliates, as.abucount)
```

```
as.incfreq
```

Transform incidence raw data to incidence frequencies (iNEXT input format)

Description

as.incfreq: transform incidence raw data (a species by sites presence-absence matrix) to incidence frequencies data (iNEXT input format, a row-sum frequencies vector contains total number of sampling units).

Usage

```
as.incfreq(x)
```

Arguments

х

a data.frame or matirx of species by sites presence-absence matrix.

bird

Value

a vector of species incidence frequencies, the first entry of the input data must be total number of sampling units.

Examples

data(ciliates)
lapply(ciliates, as.incfreq)

bird

Bird abundance-based data

Description

This data set describes 41 birds as reported by Jetz et al. (2012). It gives the two sites of species abundance to these 41 species in November 2012 at Barrington Tops National Park, Australia.

Usage

data(bird)

Format

a data frame with 41 species and two sites: North and South sites.

Source

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers A.O. (2012). The global diversity of birds in space and time. Nature, 491, 444-448.

```
library(ggplot2)
data(bird)
out <- iNEXT(bird, datatype="abundance")
ggiNEXT(out)</pre>
```

ChaoRichness

Description

ChaoRichness: estimation of species richness based on the methods proposed in Chao (1984, 1987)

Usage

```
ChaoRichness(x, datatype = "abundance", conf = 0.95)
ChaoSpecies(x, datatype = "abundance", conf = 0.95)
```

Arguments

x	a vector of species abundances or incidence frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units, followed by species incidence frequencies.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").
conf	a positive number ≤ 1 specifying the level of confidence interval.

Value

A vector of observed species richness, species richness estimate, s.e. and the associated confidence interval.

References

Chao, A. (1984) Nonparametric estimation of the number of classes in a population. Scandinavian Journal of Statistics, 11, 265-270.

Chao, A. (1987) Estimating the population size for capture-recapture data with unequal catchability. Biometrics, 43, 783-791.

See Also

ChaoShannon, ChaoSimpson

```
data(spider)
ChaoRichness(spider$Girdled, datatype="abundance")
```

ChaoShannon

Description

ChaoShannon: estimation of Shannon entropy or transformed Shannon diversity based on the method proposed by Chao et al. (2013)

Usage

```
ChaoShannon(x, datatype = "abundance", transform = FALSE,
  conf = 0.95, B = 200)
ChaoEntropy(x, datatype = "abundance", transform = FALSE,
  conf = 0.95, B = 200)
```

Arguments

X	a vector of species abundances or incidence frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units, fol- lowed by species incidence frequencies.
datatype	<pre>data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").</pre>
transform	a logical constant to compute traditional Shannon entropy index (transform=FALSE) or the transformed Shannon diversity (transform=TRUE).
conf	a positive number ≤ 1 specifying the level of confidence interval.
В	an integer specifying the number of bootstrap replications.

Value

A vector of observed Shannon entropy/diversity, estimate of entropy/diversity, s.e. and the associated confidence interval.

References

Chao, A., Wang, Y.T. & Jost, L. (2013) Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. Methods in Ecology and Evolution, 4, 1091-1100.

See Also

ChaoRichness, ChaoSimpson

```
data(spider)
ChaoShannon(spider$Girdled, datatype="abundance")
```

ChaoSimpson

Description

ChaoSimpson: estimation of Gini-Simpson index or the transformed Simpson diversity based on the methods proposed in Good (1953) and Chao et al. (2014)

Usage

```
ChaoSimpson(x, datatype = "abundance", transform = FALSE,
  conf = 0.95, B = 200)
EstSimpson(x, datatype = "abundance", transform = FALSE, conf = 0.95,
  B = 200)
```

Arguments

X	a vector of species abundances or incidence frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units, fol- lowed by species incidence frequencies.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").
transform	a logical constant to compute traditional Gini-Simpson index (transform=FALSE) or the transformed Simpson diversity (transform=TRUE).
conf	a positive number ≤ 1 specifying the level of confidence interval.
В	an integer specifying the number of bootstrap replications.

Value

a vector of observed Gini-Simpson index/diversity, index/diversity estimator, s.e. and the associated confidence interval.

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.

Good, I.J. (1953) The population frequencies of species and the estimation of population parameters. Biometrika, 40, 237-264.

See Also

ChaoRichness, ChaoShannon

ciliates

Examples

```
data(spider)
ChaoSimpson(spider$Girdled, datatype="abundance")
```

ciliates

Three soil samples

Description

A total of 51 soil samples were taken from three areas (EtoshaPan, CentralNamibDesert, SouthernNamibDesert) of Namibia. In area EtoshaPan, there were 19 soil samples and the number of observed species was 216. In area CentralNamibDesert, there were 17 soil samples and the number of observed species was 130. In area SouthernNamibDesert, there were 15 soil samples and the number of observed species species was 150.

Usage

```
data("ciliates")
```

Format

A list 3:

\$EtoshaPan is a matrix with 365 species (rows) on the following 19 sites (columns).\$CentralNamibDesert is a matrix with 365 species (rows) on the following 17 sites (columns).\$SouthernNamibDesert is a matrix with 365 species (rows) on the following 15 sites (columns).

References

Foissner, W., Agatha, S., & Berger, H. (2002) Soil Ciliates (Protozoa, Ciliophora) From Namibia (Southwest Africa), With Emphasis on Two Contrasting Environments, the Etosha Region and the Namib Desert. Denisia, 5, 1-1459.

```
data(ciliates)
out <- iNEXT(ciliates, datatype="incidence_raw")
ggiNEXT(out)</pre>
```

DataInfo

Description

DataInfo: exhibits basic data information

Usage

DataInfo(x, datatype = "abundance")

Arguments

x	a vector/matrix/list of species abundances or incidence frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units, followed by species incidence frequencies.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").

Value

a data.frame of basic data information including sample size, observed species richness, sample coverage estimate, and the first ten abundance/incidence frequency counts.

Examples

data(spider)
DataInfo(spider, datatype="abundance")

estimateD

Compute species diversity with a particular of sample size/coverage

Description

estimateD: computes species diversity (Hill numbers with q = 0, 1 and 2) with a particular userspecified level of sample size or sample coverage.

Usage

```
estimateD(x, datatype = "abundance", base = "size", level = NULL,
  conf = 0.95)
```

fortify.iNEXT

Arguments

x	a data.frame or list of species abundances or incidence frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units, followed by species incidence frequencies in each column or list.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").
base	comparison base: sample-size-based (base="size") or coverage-based (base="coverage").
level	an value specifying a particular sample size or a number (between 0 and 1) spec- ifying a particular value of sample coverage. If base="size" and level=NULL, then this function computes the diversity estimates for the minimum sample size among all sites. If base="coverage" and level=NULL, then this function com- putes the diversity estimates for the minimum sample coverage among all sites.
conf	a positive number < 1 specifying the level of confidence interval, default is 0.95. Remove C.I. by setting conf=NULL.

Value

a data.frame of species diversity table including the sample size, sample coverage, method (rarefaction or extrapolation), and diversity estimates with q = 0, 1, and 2 for the user-specified sample size or sample coverage.

Examples

```
data(spider)
estimateD(spider, "abundance", base="size", level=NULL, conf=0.95)
estimateD(spider, "abundance", base="coverage", level=NULL, conf=0.95)
data(ant)
estimateD(ant, "incidence_freq", base="coverage", level=0.985, conf=NULL)
```

fortify.iNEXT Fortify method for classes from the iNEXT package.

Description

Fortify method for classes from the iNEXT package.

Usage

```
## S3 method for class 'iNEXT'
fortify(model, data = model$iNextEst, type = 1, ...)
```

Arguments

model	iNEXT to convert into a dataframe.
data	not used by this method
type	three types of plots: sample-size-based rarefaction/extrapolation curve (type = 1); sample completeness curve (type = 2); coverage-based rarefaction/extrapolation curve (type = 3).
•••	not used by this method

Examples

```
data(spider)
# single-assemblage abundance data
out1 <- iNEXT(spider$Girdled, q=0, datatype="abundance")
ggplot2::fortify(out1, type=1)</pre>
```

```
ggiNEXT
```

ggplot2 extension for an iNEXT object

Description

ggiNEXT: the ggplot extension for iNEXT Object to plot sample-size- and coverage-based rarefaction/extrapolation curves along with a bridging sample completeness curve

Usage

```
ggiNEXT(x, type = 1, se = TRUE, facet.var = "none",
  color.var = "site", grey = FALSE)
## S3 method for class 'iNEXT'
ggiNEXT(x, type = 1, se = TRUE, facet.var = "none",
  color.var = "site", grey = FALSE)
## Default S3 method:
ggiNEXT(x, ...)
```

Arguments

х	an iNEXT object computed by iNEXT.
type	three types of plots: sample-size-based rarefaction/extrapolation curve (type = 1); sample completeness curve (type = 2); coverage-based rarefaction/extrapolation curve (type = 3).
se	a logical variable to display confidence interval around the estimated sampling curve.

iNEXT

facet.var	create a separate plot for each value of a specified variable: no separation
	(facet.var="none"); a separate plot for each diversity order (facet.var="order");
	a separate plot for each site (facet.var="site"); a separate plot for each com-
	bination of order x site (facet.var="both").
color.var	create curves in different colors for values of a specified variable: all curves are
	in the same color (color.var="none"); use different colors for diversity orders
	<pre>(color.var="order"); use different colors for sites (color.var="site"); use</pre>
	different colors for combinations of order x site (color.var="both").
grey	a logical variable to display grey and white ggplot2 theme.
	other arguments passed on to methods. Not currently used.

Value

a ggplot2 object

Examples

```
data(spider)
# single-assemblage abundance data
out1 <- iNEXT(spider$Girdled, q=0, datatype="abundance")</pre>
ggiNEXT(x=out1, type=1)
ggiNEXT(x=out1, type=2)
ggiNEXT(x=out1, type=3)
## Not run:
# single-assemblage incidence data with three orders q
data(ant)
size <- round(seq(10, 500, length.out=20))</pre>
y <- iNEXT(ant$h500m, q=c(0,1,2), datatype="incidence_freq", size=size, se=FALSE)</pre>
ggiNEXT(y, se=FALSE, color.var="order")
# multiple-assemblage abundance data with three orders q
z <- iNEXT(spider, q=c(0,1,2), datatype="abundance")</pre>
ggiNEXT(z, facet.var="site", color.var="order")
ggiNEXT(z, facet.var="both", color.var="both")
```

End(Not run)

```
iNEXT
```

iNterpolation and EXTrapolation of Hill number

Description

iNEXT: Interpolation and extrapolation of Hill number with order q

Usage

```
iNEXT(x, q = 0, datatype = "abundance", size = NULL,
endpoint = NULL, knots = 40, se = TRUE, conf = 0.95,
nboot = 50)
```

Arguments

frequencies. If datatype = "incidence", then the first entry of the input data must be total number of sampling units in each column or list.
a numeric value specifying the diversity order of Hill number .
data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence_freq") or species by sampling-units incidence matrix (datatype = "incidence_raw").
an integer vector of sample sizes (number of individuals or sampling units) for which diversity estimates will be computed. If NULL, then diversity estimates will be computed for those sample sizes determined by the specified/default endpoint and knots.
an integer specifying the sample size that is the endpoint for rarefaction/extrapolation. If NULL, then endpoint = double reference sample size.
an integer specifying the number of equally-spaced knots (say K, default is 40) between size 1 and the endpoint; each knot represents a particular sample size for which diversity estimate will be calculated. If the endpoint is smaller than the reference sample size, then iNEXT() computes only the rarefaction esimates for approximately K evenly spaced knots. If the endpoint is larger than the reference sample size, then iNEXT() computes rarefaction estimates for approximately K/2 evenly spaced knots between sample size 1 and the reference sample size, and computes extrapolation estimates for approximately K/2 evenly spaced knots between the reference sample size and the endpoint.
a logical variable to calculate the bootstrap standard error and conf confidence interval.
a positive number < 1 specifying the level of confidence interval, default is 0.95.
an integer specifying the number of replications.

Value

a list of three objects: \$DataInfo for summarizing data information; \$iNextEst for showing diversity estimates for rarefied and extrapolated samples along with related statistics; and \$AsyEst for showing asymptotic diversity estimates along with related statistics.

```
## example for abundance based data (list of vector)
data(spider)
out1 <- iNEXT(spider, q=0, datatype="abundance")
out1$DataInfo # showing basic data information.
out1$AsyEst # showing asymptotic diversity estimates.
out1$iNextEst # showing diversity estimates with rarefied and extrapolated.
## example for abundance based data (data.frame)
data(bird)
out2 <- iNEXT(bird, q=0, datatype="abundance")
ggiNEXT(out2)</pre>
```

plot.iNEXT

```
## Not run:
## example for incidence frequencies based data (list of data.frame)
data(ant)
t <- round(seq(10, 500, length.out=20))
out3 <- iNEXT(ant$h500m, q=1, datatype="incidence_freq", size=t, se=FALSE)
out3$iNextEst
```

End(Not run)

plot.iNEXT Plotting iNEXT object

Description

plot. iNEXT: Plotting method for objects inheriting from class "iNEXT"

Usage

```
## S3 method for class 'iNEXT'
plot(x, type = 1, se = TRUE, show.legend = TRUE,
   show.main = TRUE, col = NULL, ...)
```

Arguments

х	an iNEXT object computed by iNEXT.
type	three types of plots: sample-size-based rarefaction/extrapolation curve (type = 1); sample completeness curve (type = 2); coverage-based rarefaction/extrapolation curve (type = 3).
se	a logical variable to display confidence interval around the estimated sampling curve.
show.legend	a logical variable to display legend.
show.main	a logical variable to display title.
col	a vector for plotting color
	arguments to be passed to methods, such as graphical parameters (par).

```
data(spider)
# single-assemblage abundance data
out1 <- iNEXT(spider$Girdled, q=0, datatype="abundance")
plot(x=out1, type=1)
plot(x=out1, type=2)
plot(x=out1, type=3)</pre>
```

print.iNEXT

Description

print.iNEXT: Print method for objects inheriting from class "iNEXT"

Usage

S3 method for class 'iNEXT'
print(x, ...)

Arguments

х	an iNEXT object computed by iNEXT.
	additional arguments.

spider

Spider species abundances/frequencies

Description

Spider species abundances/frequencies in two canopy manipulation treatments, Girdled and Logged, of hemlock trees (Ellison et al. 2010, Sackett et al. 2011). In the Girdled treatment site, there were 26 species among 168 individuals; in the Logged treatment site, there were 37 species among 252 individuals.

Usage

data(spider)

Format

The format for each site is a list of species abundances: List of 2 \$ Girdled: num [1:26] 46 22 17 15 15 9 8 6 6 4 ... \$ Logged : num [1:37] 88 22 16 15 13 10 8 8 7 7 ...

spider

References

Ellison, A.M., Barker-Plotkin, A.A., Foster, D.R. & Orwig, D.A. (2010) Experimentally testing the role of foundation species in forests: the Harvard forest hemlock removal experiment. Methods in Ecology and Evolution, 1, 168-179.

Sackett, T.E., Record, S., Bewick, S., Baiser, B., Sanders, N.J. & Ellison, A.M. (2011) Response of macroarthropod assemblages to the loss of hemlock (Tsuga canadensis), a foundation species. Ecosphere, 2, art74.

Index

*Topic datasets ant, 3 bird, 5 ciliates, 9 spider, 16 ant, 3 as.abucount,4 as.incfreq,4 bird, 5 ChaoEntropy (ChaoShannon), 7 ChaoRichness, 6, 7, 8 ChaoShannon, 6, 7, 8 ChaoSimpson, 6, 7, 8 ChaoSpecies (ChaoRichness), 6 ciliates, 9 DataInfo, 10 estimateD, 10 $\texttt{EstSimpson} \ (\texttt{ChaoSimpson}), \ 8$ fortify.iNEXT, 11 ggiNEXT, 12 ggplot, 12 iNEXT, 12, 13, 15, 16 ${\tt iNEXT-package, 2}$ par, <u>15</u> plot.iNEXT, 15 print.iNEXT, 16

spider, 16