

Package ‘iNextPD’

March 11, 2017

Type Package

Title Interpolation and Extrapolation for Phylogenetic Diversity

Version 0.3.2

Date 2017-03-11

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

Description Interpolation and extrapolation for phylogenetic diversity.

URL <http://chao.stat.nthu.edu.tw/blog/software-download/>

License GPL (>= 3)

Depends R (>= 3.1.2)

Imports stats, graphics, ade4, ggplot2, Rcpp, iNEXT

BugReports <https://github.com/JohnsonHsieh/iNextPD/issues>

LazyLoad yes

RoxygenNote 5.0.1

Suggests knitr, rmarkdown, gridExtra, ggthemes

VignetteBuilder knitr

NeedsCompilation no

Author T. C. Hsieh [aut, cre, cph],
Chun-Huo Chiu [ths],
Anne Chao [ths]

Repository CRAN

Date/Publication 2017-03-11 09:18:39

R topics documented:

bird	2
Coverage	3
estimatePD	3
estPD	4
ExpandData	5

fortify.iNextPD	6
ggiNEXT	7
iNextPD	8
PDBoot	9
plot.iNextPD	10
print.iNextPD	11
SPBoot	12
Index	13

bird	<i>Bird phylogeny and survey data</i>
------	---------------------------------------

Description

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

Usage

bird

Format

A data frame with 53940 rows and 10 variables:

tre is a character string giving the phylogenetic tree in Newick format.

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

inic is a list of two site (data.frame) for species by sampling-units incidence matrix.

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.

Examples

```
data(bird)
bird.phy <- ade4::newick2phylog(bird$tre)
plot(bird.phy)
bird.abun <- bird$abun
bird.lab <- rownames(bird$abun)
ade4::table.phylog(bird.abun, bird.phy, csize=4, f.phylog=0.7)
```

Coverage	<i>Compute sample coverage</i>
----------	--------------------------------

Description

Coverage: Compute sample coverage

Usage

```
Coverage(x, datatype, m)
```

Arguments

x	a vector/matrix/list of species abundances/incidence frequencies or a matrix of incidence table. If <code>datatype = "incidence_freq"</code> , 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 (<code>datatype = "abundance"</code>), sampling-unit-based incidence frequencies data (<code>datatype = "incidence_freq"</code>) or species by sampling-units incidence matrix (<code>datatype = "incidence_raw"</code>).
m	an integer vector of sample sizes (number of individuals or sampling units) for which diversity estimates will be computed.

Value

a data.frame with sample size and sample coverage.

Examples

```
data(bird)
Coverage(bird$abun, datatype="abundance", m=c(10,50,100,150,200))
```

estimatePD	<i>Compute phylogenetic diversity with a particular of sample size/coverage</i>
------------	---

Description

estimatePD: computes phylogenetic diversity (Hill numbers with $q = 0, 1$ and 2) with a particular user-specified level of sample size or sample coverage.

Usage

```
estimatePD(x, labels, phy, datatype = "abundance", base = "size",
  level = NULL, conf = 0.95, digits = 4)
```

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.
labels	species names for object x
phy	a phylog object for input phylo-tree
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), 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 integer specifying a particular sample size or a number (between 0 and 1) specifying 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 computes 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.
digits	integer indicating the number of decimal places round to be used.

Value

a data.frame of phylogenetic 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(bird)
bird.abu <- bird$abun
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
estimatePD(bird.abu, bird.lab, bird.phy, "abundance", base="size", level=NULL, conf=NULL)
## Not run:
estimatePD(bird.abu, bird.lab, bird.phy, "abundance", base="size", level=NULL)

## End(Not run)
```

 estPD

Estimate asymptotic phylogenetic diversity

Description

estPD: Estimate asymptotic phylogenetic diversity with order $q = 0, 1, 2$.

Usage

```
estPD(x, labels, phy, q = 0, datatype = "abundance", se = FALSE,
      conf = 0.95)
```

Arguments

x a vector/matrix/list of species abundances or a matrix of raw incidence table.

labels a vector of species name for input data.

phy a phylogenetic tree with "phylog" class.

q a numeric value specifying the diversity order of Hill number.

datatype data type of input data: individual-based abundance data (datatype = "abundance"), species by sampling-units incidence matrix (datatype = "incidence_raw").

se a logical variable to calculate the bootstrap standard error and conf confidence interval.

conf a positive number < 1 specifying the level of confidence interval, default is 0.95.

Value

a data.frame with sample size and sample coverage.

Examples

```
# abundance-based example
data(bird)
bird.abu <- bird$abun
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
estPD(bird.abu, labels=bird.lab, phy=bird.phy, q=0, datatype="abundance")

# incidence_based example
bird.inc <- bird$inci
estPD(bird.inc, labels=bird.lab, phy=bird.phy, q=0, datatype="incidence_raw", se=TRUE)
```

ExpandData

Expand branch abundance/incience and branch length

Description

ExpandData: Expand branch abundance/incience and branch length

Usage

```
ExpandData(x, labels, phy, datatype = "abundance")
```

Arguments

x	a vector/matrix/list of species abundances or a matrix of raw incidence table.
labels	a vector of species name for input data.
phy	a phylogenetic tree with "phylog" class.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), species by sampling-units incidence matrix (datatype = "incidence_raw").

Value

a data.frame with sample size and sample coverage.

Examples

```
data(bird)
bird.abu <- bird$abun
bird.inc <- bird$inci
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
ExpandData(bird.abu, labels=bird.lab, phy=bird.phy, datatype="abundance")
ExpandData(bird.inc, labels=bird.lab, phy=bird.phy, datatype="incidence_raw")
```

fortify.iNextPD

Fortify method for classes from the iNextPD package.

Description

Fortify method for classes from the iNextPD package.

Usage

```
## S3 method for class 'iNextPD'
fortify(model, data = model$iNextPDEst, type = 1,
        se = TRUE, ...)
```

Arguments

model	iNextPD 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).
se	a logical variable to calculate the bootstrap standard error and conf confidence interval.
...	not used by this method

Examples

```
# single-assembly abundance data
data(bird)
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
out1 <- iNextPD(bird$abun$North.site, bird.lab, bird.phy,
  q=0, datatype="abundance", endpoint=400)
ggplot2::fortify(out1, type=1)
```

ggiNEXT

*ggplot2 extension for an iNextPD object***Description**

ggiNEXT: the ggplot extension for `iNextPD` 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 'iNextPD'
```

```
ggiNEXT(x, type = 1, se = TRUE, facet.var = "none",
  color.var = "site", grey = FALSE)
```

```
## Default S3 method:
```

```
ggiNEXT(x, ...)
```

Arguments

<code>x</code>	an <code>iNextPD</code> object computed by <code>iNextPD</code> .
<code>type</code>	three types of plots: sample-size-based rarefaction/extrapolation curve (<code>type = 1</code>); sample completeness curve (<code>type = 2</code>); coverage-based rarefaction/extrapolation curve (<code>type = 3</code>).
<code>se</code>	a logical variable to display confidence interval around the estimated sampling curve.
<code>facet.var</code>	create a separate plot for each value of a specified variable: no separation (<code>facet.var="none"</code>); a separate plot for each diversity order (<code>facet.var="order"</code>); a separate plot for each site (<code>facet.var="site"</code>); a separate plot for each combination of order x site (<code>facet.var="both"</code>).
<code>color.var</code>	create curves in different colors for values of a specified variable: all curves are in the same color (<code>color.var="none"</code>); use different colors for diversity orders (<code>color.var="order"</code>); use different colors for sites (<code>color.var="site"</code>); use different colors for combinations of order x site (<code>color.var="both"</code>).
<code>grey</code>	a logical variable to display grey and white ggplot2 theme.
<code>...</code>	other arguments passed on to methods. Not currently used.

Value

a ggplot2 object

Examples

```
# single-assemblage abundance data
data(bird)
bird.phy <- ade4::newick2phylog(bird$tre)
bird.lab <- rownames(bird$abun)
out1 <- iNextPD(bird$abun$North.site, bird.lab, bird.phy,
               q=1, datatype="abundance", endpoint=400, se=TRUE)
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
out2 <- iNextPD(bird$inci$North.site, bird.lab, bird.phy,
               q=c(0,1,2), datatype="incidence_raw", endpoint=25)
ggiNEXT(out2, se=FALSE, color.var="order")

# multiple-assemblage abundance data with three orders q
out3 <- iNextPD(bird$abun, bird.lab, bird.phy,
               q=c(0,1,2), datatype="abundance", endpoint=400)
ggiNEXT(out3, facet.var="site", color.var="order")
ggiNEXT(out3, facet.var="both", color.var="both")

## End(Not run)
```

iNextPD

iNterpolation and EXTrapolation of Hill number

Description

iNextPD: Interpolation and extrapolation of Hill number with order q .

Usage

```
iNextPD(x, labels, phy, q = 0, datatype = "abundance", size = NULL,
        endpoint = NULL, knots = 40, se = FALSE, conf = 0.95, nboot = 50)
```

Arguments

<code>x</code>	a matrix, data.frame (species by sites), or list of species abundances or incidence data.
<code>labels</code>	species names for object <code>x</code> .
<code>phy</code>	a phylog object for input phylo-tree.
<code>q</code>	a numeric value specifying the diversity order of Hill number .

datatype	data type of input data: individual-based abundance data (datatype = "abundance"), or species by sampling-units incidence matrix (datatype = "incidence_raw").
size	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 .
endpoint	an integer specifying the sample size that is the endpoint for rarefaction/extrapolation. If NULL, then endpoint = double reference sample size.
knots	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 iNextPD() computes only the rarefaction estimates for approximately K evenly spaced knots. If the endpoint is larger than the reference sample size, then iNextPD() 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.
se	a logical variable to calculate the bootstrap standard error and conf confidence interval.
conf	a positive number < 1 specifying the level of confidence interval, default is 0.95.
nboot	an integer specifying the number of replications.

Value

a list of three objects: \$DataInfo for summarizing data information; \$iNextPDEst for showing diversity estimates for rarefied and extrapolated samples along with related statistics; \$AsyPDEst for showing asymptotic diversity estimates along with related statistics, and \$ExpandData (xi, Li, i=1,2,...,B).

Examples

```
data(bird)
bird.abu <- bird$abun
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
iNextPD(bird.abu, labels=bird.lab, phy=bird.phy, q=0, datatype="abundance")
```

PDBoot

Expand bootstrapping branch abundance/incience and branch length

Description

PDBoot: Expand bootstrapping branch abundance/incience and branch length

Usage

```
PDBoot(x, labels, phy, datatype = "abundance")
```

Arguments

x	a vector/matrix/list of species abundances or a matrix of raw incidence table.
labels	a vector of species name for input data.
phy	a phylogenetic tree with "phylog" class.
datatype	data type of input data: individual-based abundance data (datatype = "abundance"), species by sampling-units incidence matrix (datatype = "incidence_raw").

Value

a list of data.frame with bootstrapping branch abundance/incience and branch length.

Examples

```
data(bird)
bird.lab <- rownames(bird$abun)
bird.phy <- ade4::newick2phylog(bird$tre)
bird.inc <- bird$inci
PDBoot(bird.inc, labels=bird.lab, phy=bird.phy, datatype="incidence_raw")
```

plot.iNextPD

Plotting iNextPD object

Description

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

Usage

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

Arguments

x	an iNextPD object computed by iNextPD .
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.
xlab	a title for the x axis.
ylab	a title for the y axis.
...	arguments to be passed to methods, such as graphical parameters (par).

Examples

```
# single-assemblage abundance data
data(bird)
bird.phy <- ade4::newick2phylog(bird$tre)
bird.lab <- rownames(bird$abun)
out1 <- iNextPD(bird$abun$North.site, bird.lab, bird.phy,
               q=1, datatype="abundance", endpoint=500)
plot(x=out1, type=1)
plot(x=out1, type=2)
plot(x=out1, type=3)
```

print.iNextPD	<i>Printing iNextPD object</i>
---------------	--------------------------------

Description

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

Usage

```
## S3 method for class 'iNextPD'
print(x, ...)
```

Arguments

x	an iNextPD object computed by iNextPD .
...	additional arguments.

SPBoot	<i>Estimation of species relative abundance or detection probability distribution</i>
--------	---

Description

SPBoot: Expand bootstrapping species relative abundance or detection probability

Usage

```
SPBoot(x, datatype = "abundance")
```

Arguments

x	a vector/matrix/list of species abundances or a matrix of raw incidence table.
datatype	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 list of vector with species relative abundance or detection probability distribution.

Examples

```
data(bird)
bird.inc <- bird$inci
SPBoot(bird$abun, datatype="abundance")
SPBoot(bird$inci, datatype="incidence_raw")
```

Index

*Topic **datasets**

bird, [2](#)

bird, [2](#)

Coverage, [3](#)

estimatePD, [3](#)

estPD, [4](#)

ExpandData, [5](#)

fortify.iNextPD, [6](#)

ggiNEXT, [7](#)

iNextPD, [7](#), [8](#), [10](#), [11](#)

par, [11](#)

PDBoot, [9](#)

plot.iNextPD, [10](#)

print.iNextPD, [11](#)

SPBoot, [12](#)