

# Package ‘iNEXT’

March 24, 2024

**Type** Package

**Title** Interpolation and Extrapolation for Species Diversity

**Version** 3.0.1

**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 curves for species diversity (Hill numbers) based on individual-based abundance data or sampling-unit-based incidence data; see Chao and others (2014, Ecological Monographs) for pertinent theory and methodologies, and Hsieh, Ma and Chao (2016, Methods in Ecology and Evolution) for an introduction of the R package.

**License** GPL ( $\geq 3$ )

**URL** [http://chao.stat.nthu.edu.tw/wordpress/software\\_download/](http://chao.stat.nthu.edu.tw/wordpress/software_download/)

**Maintainer** T. C. Hsieh <[euler96@gmail.com](mailto:euler96@gmail.com)>

**Depends** R ( $\geq 3.4$ )

**Imports** stats, graphics, ggplot2, reshape2, Rcpp

**BugReports** <https://github.com/JohnsonHsieh/iNEXT/issues>

**LazyLoad** yes

**RoxygenNote** 7.2.0

**Suggests** testthat, knitr, rmarkdown, gridExtra, ggthemes

**VignetteBuilder** knitr

**ByteCompile** true

**LinkingTo** Rcpp

**NeedsCompilation** yes

**Repository** CRAN

**Date/Publication** 2024-03-24 17:10:09 UTC

## R topics documented:

iNEXT-package . . . . .	2
ant . . . . .	3
as.abucount . . . . .	4
as.incfreq . . . . .	4
bird . . . . .	5
ChaoRichness . . . . .	6
ChaoShannon . . . . .	7
ChaoSimpson . . . . .	8
ciliates . . . . .	9
DataInfo . . . . .	9
estimateD . . . . .	10
fortify.iNEXT . . . . .	12
ggiNEXT . . . . .	12
iNEXT . . . . .	14
plot.iNEXT . . . . .	16
print.iNEXT . . . . .	17
spider . . . . .	17
<b>Index</b>	<b>19</b>

---

iNEXT-package

*Interpolation and extrapolation for species diversity*

---

## Description

iNEXT (iNterpolation and EXTrapolation) provides functions to compute and plot two types (sample-size- 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/sites. 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. A short review of the theoretical background and a brief description of methods are included in an application paper by Hsieh, Ma & Chao (2016). An online version (<https://chao.shinyapps.io/iNEXTOnline/>) is also available for users without an R background.



## 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	<i>Transform abundance raw data to abundance row-sum counts (iNEXT input format)</i>
-------------	--

---

## Description

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

## Usage

```
as.abucount(x)
```

## Arguments

x a data.frame or matrix (species by sites matrix).

## Value

a vector of species abundance row-sum counts.

## Examples

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

---

as.incfreq	<i>Transform incidence raw data to incidence frequencies (iNEXT input format)</i>
------------	---

---

## Description

as.incfreq: transform incidence raw data (a species by sites detection/non-detection or presence/absence matrix) to incidence frequencies data (iNEXT input format): the first element is the total number of sampling units, followed by the vector of species frequencies. Here species frequencies represent the row sums of the incidence raw matrix.

## Usage

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

**Arguments**

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

**Value**

a vector of species incidence frequencies, the first element is the total number of sampling units.

**Examples**

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

---

bird	<i>Bird data (datatype = "abundance")</i>
------	---

---

**Description**

This data set includes the abundances of 41 bird species collected in two sites (the North and South sites) at the Barrington Tops National Park, Australia (Chao et al. 2015)

**Usage**

```
data(bird)
```

**Format**

a data.frame with 41 species (rows) and two sites (columns).

**Source**

Chao, A., Chiu, C.-H., Hsieh, T. C., Davis, T., Nipperess, D., and Faith, D. (2015) Rarefaction and extrapolation of phylogenetic diversity. *Methods in Ecology and Evolution*, 6, 380-388.

**Examples**

```
data(bird)
## Not run:
out <- iNEXT(bird, datatype="abundance")
ggiNEXT(out)

## End(Not run)
```

---

ChaoRichness

*Estimation of species richness*


---

### Description

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

### Usage

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

### Arguments

x	a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If datatype = "incidence_freq", 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 $\leq 1$ specifying the level of confidence interval.

### Value

A data.frame 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](#)

### Examples

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

---

ChaoShannon

*Estimation of Shannon entropy/diversity*

---

### 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)
```

### Arguments

x	a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If datatype = "incidence_freq", 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").
transform	a logical constant to compute traditional Shannon entropy index (transform=FALSE) or the transformed Shannon diversity (transform=TRUE).
conf	a positive number $\leq 1$ specifying the level of confidence interval.
B	an integer specifying the number of bootstrap replications.

### Value

A data.frame 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](#)

### Examples

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

ChaoSimpson

*Estimation of Gini-Simpson index or Simpson diversity***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)
```

**Arguments**

x	a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If datatype = "incidence_freq", 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").
transform	a logical constant to compute traditional Gini-Simpson index (transform=FALSE) or the transformed Simpson diversity (transform=TRUE).
conf	a positive number $\leq 1$ specifying the level of confidence interval.
B	an integer specifying the number of bootstrap replications.

**Value**

a data.frame 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](#)

**Examples**

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



---

`ciliates`*Ciliates data (datatype = "incidence\_raw")*

---

**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 was 150. The total number of species in the three areas was 365. The data include three lists of matrices; each matrix is a species by soil-sample matrix ("1" for a detection, and "0" otherwise).

**Usage**

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

**Format**

A list 3 matrices:

\$EtoshaPan is a matrix with 365 species (rows) and 19 soil samples (columns).

\$CentralNamibDesert is a matrix with 365 species (rows) and 17 soil samples (columns).

\$SouthernNamibDesert is a matrix with 365 species (rows) and 15 soil samples (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.

**Examples**

```
data(ciliates)
## Not run:
out <- iNEXT(ciliates, datatype = "incidence_raw")
ggiNEXT(out)

## End(Not run)
```

---

`DataInfo`*Exhibit basic data information*

---

**Description**

DataInfo: exhibits basic data information

**Usage**

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

**Arguments**

**x** a vector/matrix/data.frame/list of species abundances or incidence frequencies. If `datatype = "incidence_freq"`, 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	<i>Compute species diversity with a particular level of sample size/coverage</i>
-----------	--

---

**Description**

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

**Usage**

```
estimateD(
  x,
  q = c(0, 1, 2),
  datatype = "abundance",
  base = "size",
  level = NULL,
  nboot = 50,
  conf = 0.95
)
```

**Arguments**

x	a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If datatype = "incidence_freq", 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.
q	a number or vector specifying the diversity order(s) of Hill numbers.
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	a sequence specifying the particular sample sizes or sample coverages (between 0 and 1). If base="size" and level=NULL, then this function computes the diversity estimates for the minimum among all double reference sample sizes. If base="coverage" and level=NULL, then this function computes the diversity estimates for the minimum among all the coverage values for samples extrapolated to double the reference sample sizes.
nboot	the number of bootstrap times to obtain confidence interval. If confidence interval is not desired, use 0 to skip this time-consuming step; default is 50.
conf	a positive number < 1 specifying the level of confidence interval; default is 0.95.

**Value**

a data.frame of species diversity table including the sample size, sample coverage, method (rarefaction or extrapolation), and diversity estimates with the user-specified diversity orders (q values) and specified sample size or sample coverage.

**Examples**

```
data(spider)
out1 <- estimateD(spider, q = c(0,1,2), datatype = "abundance", base="size")
out1
## Not run:
out2 <- estimateD(spider, q = c(0,1,2), datatype = "abundance", base="coverage")
out2

data(ant)
out <- estimateD(ant, q = c(0,1,2), datatype = "incidence_freq", base="coverage",
                level=0.985, conf=0.95)

out

## End(Not run)
```

---

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)
```

---

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 = "Assemblage",
  grey = FALSE
```

```

)

## S3 method for class 'iNEXT'
ggiNEXT(
  x,
  type = 1,
  se = TRUE,
  facet.var = "None",
  color.var = "Assemblage",
  grey = FALSE
)

## Default S3 method:
ggiNEXT(x, ...)

```

### Arguments

<code>x</code>	an iNEXT object computed by <code>iNEXT</code> .
<code>type</code>	three types of plots: sample-size-based rarefaction/extrapolation curve (type = 1); sample completeness curve (type = 2); coverage-based rarefaction/extrapolation curve (type = 3).
<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.q"</code> ); a separate plot for each assemblage ( <code>facet.var="Assemblage"</code> ); a separate plot for each combination of order x assemblage ( <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.q"</code> ); use different colors for sites ( <code>color.var="Assemblage"</code> ); use different colors for combinations of order x assemblage ( <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(spider)
out1 <- iNEXT(spider$Girdled, q=0, datatype="abundance")
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))
y <- iNEXT(ant$h500m, q=c(0,1,2), datatype="incidence_freq", size=size, se=FALSE)
ggiNEXT(y, se=FALSE, color.var="Order.q")

# multiple-assemblage abundance data with three orders q
z <- iNEXT(spider, q=c(0,1,2), datatype="abundance")
ggiNEXT(z, facet.var="Assemblage", color.var="Order.q")
ggiNEXT(z, facet.var="Both", color.var="Both")

## End(Not run)

```

---

iNEXT

*iNterpolation and EXTrapolation of Hill numbers*


---

## 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

<code>x</code>	a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If <code>datatype = "incidence_freq"</code> , then the first entry of the input data must be total number of sampling units in each column or list.
<code>q</code>	a number or vector specifying the diversity order(s) of Hill numbers.
<code>datatype</code>	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> ).
<code>size</code>	an integer vector of sample sizes (number of individuals or sampling units) for which diversity estimates will be computed. If <code>NULL</code> , 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 the 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 iNEXT() computes only the rarefaction estimates 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.
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; default is 50.

### 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.

NOTE: From version 3.0.0, \$iNextEst has been expanded to include \$size\_based and \$coverage\_based to provide two types of confidence intervals.

### Examples

```
## Not run:
## example for abundance based data (list of vector)
data(spider)
out1 <- iNEXT(spider, q=c(0,1,2), datatype="abundance")
out1$DataInfo # showing basic data information.
out1$AsyEst # showing asymptotic diversity estimates.
out1$iNextEst$size_based
# showing diversity estimates with rarefied and extrapolated samples;
# confidence limits are obtained for fixed sample size.

out1$iNextEst$coverage_based
# showing diversity estimates with rarefied and extrapolated samples;
# confidence limits are obtained for fixed sample coverage.

## End(Not run)
## example for abundance based data (data.frame)
data(bird)
out2 <- iNEXT(bird, q=0, datatype="abundance")
out2

## 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

x	an iNEXT object computed by <a href="#">iNEXT</a> .
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 ( <a href="#">par</a> ).

### Examples

```

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)
```

---

```
print.iNEXT          Printing iNEXT object
```

---

### Description

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

### Usage

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

### Arguments

x                    an iNEXT object computed by [iNEXT](#).  
 ...                  additional arguments.

---

```
spider              Spider data (datatype = "abundance")
```

---

### Description

The data include spider species abundances 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:  
 A list of 2 vectors  
 \$ 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 ...

**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

## \* datasets

- ant, [3](#)
- bird, [5](#)
- ciliates, [9](#)
- spider, [17](#)

- ant, [3](#)
- as.abucount, [4](#)
- as.incfreq, [4](#)

- bird, [5](#)

- ChaoRichness, [6](#), [7](#), [8](#)
- ChaoShannon, [6](#), [7](#), [8](#)
- ChaoSimpson, [6](#), [7](#), [8](#)
- ciliates, [9](#)

- DataInfo, [9](#)

- estimateD, [10](#)

- fortify.iNEXT, [12](#)

- ggiNEXT, [12](#)
- ggplot, [12](#)

- iNEXT, [12](#), [13](#), [14](#), [16](#), [17](#)
- iNEXT-package, [2](#)

- par, [16](#)
- plot.iNEXT, [16](#)
- print.iNEXT, [17](#)

- spider, [17](#)