

# Package ‘sprex’

February 24, 2015

**Type** Package

**Title** Calculate Species Richness and Extrapolation Metrics

**Description** Calculate species richness functions for rarefaction and extrapolation.

**Version** 1.0

**Date** 2015-02-23

**Author** Eric Archer

**Maintainer** Eric Archer <eric.archer@noaa.gov>

**License** GNU General Public License

## R topics documented:

sprex-package . . . . .	1
ACE . . . . .	2
bootstrap.assemblage . . . . .	3
discovery.curve . . . . .	4
expand.freqs . . . . .	5
expected.num.species . . . . .	6
f.stats . . . . .	7
num.samples.required . . . . .	7
osa.old.growth . . . . .	8
osa.second.growth . . . . .	8
plot.discovery.curve . . . . .	9
sample.to.species.freq . . . . .	9
species.to.sample.freq . . . . .	10
<b>Index</b>	<b>12</b>

---

sprex-package	<i>Calculate species richness and extrapolation metrics</i>
---------------	---

---

## Description

Calculate species richness and expected number of species primarily based on algorithms in Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

Details

sprex

---

ACE	<i>Number of Unobserved Species</i>
-----	-------------------------------------

---

Description

Calculate the number of unobserved species (f0).

Usage

ACE(f)  
  
Chao1(f)  
  
Swor1(f, N)  
  
iChao1(f)  
  
jack1(f)  
  
jack2(f)

Arguments

f	a vector of species frequencies where f[i] is the number of species represented by only i samples.
N	population size.

Value

All functions return a vector containing the estimated number of species (s.est), unobserved species (f0), observed species (s.obs), and the total number of samples (n). Swor1 also returns the standard deviation of s.est as sd.s.est.

Author(s)

Eric Archer <eric.archer@noaa.gov>

References

Chao1,ACE: Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

jack1,jack2: Burnham, KP and WS Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65(3):625-633.

Swor1: Chao, A. and C.-W. Lin. 2012. Nonparametric lower bounds for species richness and shared species richness under sampling without replacement. *Biometrics* 68:912-921.

iChao1: Chiu, C-H, Wang, Y-T, Walther, BA, and A Chao. 2014. An improved nonparametric lower bound of species richness via a modified Good-Turing frequency formula. *Biometrics* 70(3):671-682.

## Examples

```
data(osa.second.growth)
f <- expand.freqs(osa.second.growth)

ace.est <- ACE(f)
chao1.est <- Chao1(f)
jack1.est <- jack1(f)
jack2.est <- jack2(f)
swor1.est <- Swor1(f, 20000)
ichao1.est <- iChao1(f)

f0.est <- cbind(
  ACE = ace.est["f0"],
  Chao1 = chao1.est["f0"],
  jack1 = jack1.est["f0"],
  jack2 = jack2.est["f0"],
  Swor1 = swor1.est["f0"],
  iChao1 = ichao1.est["f0"]
)
f0.est
```

---

bootstrap.assemblage    *Bootstrap Assemblage of Species*

---

## Description

Create bootstrap assemblage of species.

## Usage

```
bootstrap.assemblage(f, f0.func = Chao1, n.boot = 500, ...)
```

## Arguments

f	a vector of species frequencies where $f[i]$ is the number of species represented by only $i$ samples.
f0.func	function calculating the unobserved number of species ( $f_0$ ).
n.boot	number of bootstrap replicates.
...	other arguments to f0.func.

## Value

a list of bootstrap replicates of species frequencies.

## Author(s)

Eric Archer <eric.archer@noaa.gov>

## References

Chao, A., N.J. Gotelli, T.C. Hsieh, E.L. Sander, K.H. Ma, R.K. Colwell, and A.M. Ellison. 2014. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* 84(1):45-67.

---

discovery.curve

*Discovery Curve*

---

## Description

Calculate the components of a species discovery curve.

## Usage

```
discovery.curve(f, max.x = sum(f * 1:length(f)), n.pts = 100, ci = 0.95,
  f0.func = Chao1, ...)
```

## Arguments

<code>f</code>	a vector of species frequencies where <code>f[i]</code> is the number of species represented by only <code>i</code> samples.
<code>max.x</code>	the maximum number of samples to calculate the curve for. Defaults to the sample size of <code>f</code> .
<code>n.pts</code>	number of points between 0 and <code>max.x</code> to estimate.
<code>ci</code>	size of the confidence interval (0.5:1).
<code>f0.func</code>	function to use to calculate <code>f0</code> . Can be <a href="#">Chao1</a> , <a href="#">ACE</a> , <a href="#">jack1</a> , <a href="#">jack2</a> , <a href="#">iChao1</a> , or <a href="#">Swor1</a> .
<code>...</code>	other arguments to <code>f0.func</code> .

## Value

a list with:

<code>f.stats</code>	a named vector from <code>f0.func</code> .
<code>s.ind</code>	a matrix of <code>S.ind</code> estimates for each value of <code>m</code> along with the standard deviation of <code>S.ind</code> .
<code>s.ind.ci</code>	a matrix of the upper and lower confidence intervals of <code>S.ind</code> .
<code>ci.poly</code>	a matrix of points describing the <code>ci</code> polygon.
<code>rarefact.line</code>	a matrix of points defining the rarefaction line ( $\leq S_{\text{obs}}$ ).
<code>extrap.line</code>	a matrix of points defining the extrapolation line ( $> S_{\text{obs}}$ ).

## Author(s)

Eric Archer <eric.archer@noaa.gov>

## References

Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

**See Also**[plot.discovery.curve](#)**Examples**

```
data(osa.old.growth)
f <- expand.freqs(osa.old.growth)
d <- discovery.curve(f, max.x = 1200)
plot(d)
```

---

expand.freqs*Expand Frequency Matrix*

---

**Description**

Expand a matrix or data.frame of species frequencies to full vector.

**Usage**

```
expand.freqs(freq.mat)
```

**Arguments**

freq.mat	a two column matrix or data.frame where the first column is the number of samples, and the second column is the number of species represented by with that many samples.
----------	--

**Value**

a vector(f) of species frequencies where each element (f[i]) is the number of species represented by only i samples.

**Author(s)**

Eric Archer <eric.archer@noaa.gov>

**Examples**

```
data(osa.old.growth)
f <- expand.freqs(osa.old.growth)
f
```

---

expected.num.species	<i>Expected Number of Species</i>
----------------------	-----------------------------------

---

## Description

Calculate the expected number of species for a given sample size.

## Usage

```
expected.num.species(m, f, f0.func = Chao1, ...)
```

## Arguments

<code>m</code>	number of samples.
<code>f</code>	a vector of species frequencies where <code>f[i]</code> is the number of species represented by only <code>i</code> samples.
<code>f0.func</code>	a function that computes the number of unobserved species ( <code>f0</code> ).
<code>...</code>	other arguments to <code>f0.func</code> .

## Value

a vector of the estimated number of species (`s.ind`) seen in `m` samples, and the standard deviation (`sd.s.ind`).

## Author(s)

Eric Archer <eric.archer@noaa.gov>

## References

Eqns 4, 5, 9, and 10 in Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

## Examples

```
data(osa.old.growth)
f <- expand.freqs(osa.old.growth)
expected.num.species(60, f)
```

---

f.stats	<i>Frequency Vector Statistics</i>
---------	------------------------------------

---

**Description**

Number of observed species and samples in species frequency vector.

**Usage**

```
f.stats(f)
```

**Arguments**

f	a vector of species frequencies where f[i] is the number of species represented by only i samples.
---	--

**Value**

a vector of the number of observed species (s.obs), and the total number of samples (n).

**Author(s)**

Eric Archer <eric.archer@noaa.gov>

**Examples**

```
data(osa.second.growth)
f <- expand.freqs(osa.second.growth)
f.stats(f)
```

---

num.samples.required	<i>Number of Samples Required</i>
----------------------	-----------------------------------

---

**Description**

Calculate the number of samples to required to observe a given proportion of the total number of species.

**Usage**

```
num.samples.required(g, f, f0.func = Chao1)
```

**Arguments**

g	propotion of total number of species.
f	a vector of species frequencies where f[i] is the number of species represented by only i samples.
f0.func	a function that computes the number of unobserved species (f0).

**Value**

a vector containing of the estimated number of samples (m.g) required to observe g percent of the total number of species.

**Author(s)**

Eric Archer <eric.archer@noaa.gov>

**References**

Eqn 11 in Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

**Examples**

```
data(osa.old.growth)
f <- expand.freqs(osa.old.growth)
num.samples.required(0.6, f)
```

---

osa.old.growth	<i>Osa old growth</i>
----------------	-----------------------

---

**Description**

Osa old growth

**Usage**

```
data(osa.old.growth)
```

**References**

Janzen DH (1973) Sweep samples of tropical foliage insects: effects of seasons, vegetation types, elevation, time of day, and insularity. *Ecology* 54:687-708.

---

osa.second.growth	<i>Osa second growth</i>
-------------------	--------------------------

---

**Description**

Osa second growth

**Usage**

```
data(osa.second.growth)
```

**References**

Janzen DH (1973) Sweep samples of tropical foliage insects: description of study sites, with data on species abundances and size distributions. *Ecology* 54:659-86.



---

plot.discovery.curve    *Plot a Discovery Curve*


---

### Description

Plot a species discovery curve.

### Usage

```
## S3 method for class 'discovery.curve'
plot(x, col = "darksalmon", lwd = 2,
     xlab = "# Samples", ylab = "n", add = FALSE, ...)
```

### Arguments

x	result of a call to discovery.curve.
col	color of confidence interval polygon and line denoting s.est.
lwd	line widths.
xlab,ylab	labels of x and y axes. Only used if add is TRUE.
add	logical. If TRUE, polygon and lines are added to the current plot.
...	other arguments passed to plot (ignored).

### Author(s)

Eric Archer <eric.archer@noaa.gov>

### References

Colwell, R.K., A. Chao, N.J. Gotelli, S.-Y. Lin, C.X. Mao, R.L. Chazdon, and J.T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5(1):3-21.

### See Also

[discovery.curve](#)

---

sample.to.species.freq

*Create Vector of Species Frequencies*

---

### Description

Create vector of species frequencies from vector of sample frequencies.

### Usage

```
sample.to.species.freq(x)
```

**Arguments**

`x` a vector where `x[i]` is of the number of samples in the *i*-th species.

**Value**

a vector(`f`) of species frequencies where `f[i]` is the number of species represented by only *i* samples.

**Author(s)**

Eric Archer <eric.archer@noaa.gov>

**See Also**

species.to.sample.freq

**Examples**

```
x <- sample(1:100, 30, rep = TRUE)
f <- sample.to.species.freq(x)
print(x)
print(f)
```

---

species.to.sample.freq

*Create Vector of Sample Frequencies*

---

**Description**

Create vector of sample frequencies from vector of species frequencies.

**Usage**

```
species.to.sample.freq(f)
```

**Arguments**

`f` a vector of species frequencies where `f[i]` is the number of species represented by only *i* samples.

**Value**

a vector(`x`) where `x[i]` is of the number of samples in the *i*-th species.

**Author(s)**

Eric Archer <eric.archer@noaa.gov>

**See Also**

sample.to.species.freq

**Examples**

```
data(osa.old.growth)
f <- expand.freqs(osa.old.growth)
x <- species.to.sample.freq(f)
print(f)
print(x)
```

# Index

## \*Topic **datasets**

osa.old.growth, [8](#)

osa.second.growth, [8](#)

## \*Topic **package**

sprex-package, [1](#)

ACE, [2](#), [4](#)

bootstrap.assemblage, [3](#)

Chao1, [4](#)

Chao1 (ACE), [2](#)

discovery.curve, [4](#), [9](#)

expand.freqs, [5](#)

expected.num.species, [6](#)

f.stats, [7](#)

f0 (ACE), [2](#)

iChao1, [4](#)

iChao1 (ACE), [2](#)

jack1, [4](#)

jack1 (ACE), [2](#)

jack2, [4](#)

jack2 (ACE), [2](#)

num.samples.required, [7](#)

osa.old.growth, [8](#)

osa.second.growth, [8](#)

plot.discovery.curve, [5](#), [9](#)

sample.to.species.freq, [9](#)

species.to.sample.freq, [10](#)

sprex (sprex-package), [1](#)

sprex-package, [1](#)

Swor1, [4](#)

Swor1 (ACE), [2](#)