Package ‘simboot’

February 20, 2015

Type Package
Title Simultaneous inference for diversity indices.
Version 0.2-5
Date 2014-04-10
Author Ralph Scherer, Philip Pullmann
Maintainer Ralph Scherer <shearer.ra76@gmail.com>
Description Package simboot provides estimation of simultaneous bootstrap and asymptotic confidence intervals for diversity indices, namely the Shannon and the Simpson index. Several pre-specified multiple comparison types are available to choose. Further user-defined contrast matrices are applicable. In addition, simboot estimates adjusted as well as unadjusted p-values for two of the three proposed bootstrap methods. Further simboot allows for comparing biological diversities of two or more groups while simultaneously testing a user-defined selection of Hill numbers of orders q, which are considered as appropriate and useful indices for measuring diversity.
License GPL (>= 2)
URL https://github.com/shearer/simboot,
http://shearer.github.io/simboot/

BugReports https://github.com/shearer/simboot/issues
Depends boot, mvtnorm
LazyLoad yes
NeedsCompilation no
Repository CRAN
Date/Publication 2014-04-16 00:28:09

R topics documented:

  simboot-package ...................................................... 2
  asht ................................................................. 3
  Bacteria .............................................................. 4
  Bouterp ............................................................... 6
Simultaneous inference for diversity indices.

Description

Package simboot provides estimation of simultaneous bootstrap and asymptotic confidence intervals for diversity indices, namely the Shannon and the Simpson index. Several pre-specified multiple-comparison types are available. Further user-defined contrast matrices are applicable. In addition, simboot estimates adjusted as well as unadjusted \( p \)-values for two of the three proposed bootstrap methods. Further simboot allows for comparing biological diversities of two or more groups with simultaneously testing a user-defined selection of Hill numbers of orders \( q \), which are considered appropriate and useful indices for measuring diversity.

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>simboot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>0.2-5</td>
</tr>
<tr>
<td>Date:</td>
<td>2014-04-10</td>
</tr>
<tr>
<td>License:</td>
<td>GPL (&gt;= 2)</td>
</tr>
<tr>
<td>LazyLoad:</td>
<td>yes</td>
</tr>
</tbody>
</table>
Author(s)
Ralph Scherer\ Philip Pallmann\ Maintainer: Ralph Scherer <shearer.ra76@gmail.com>

References
Scherer, R. and Schaarschmidt, F. (2013) Simultaneous confidence intervals for comparing biodi-
versity indices estimated from overdispersed count data. Biometrical Journal 55, 246–263.
Evaluation of the methods in sbdiv
Pallmann, P. et al. (2012) Assessing group differences in biodiversity by simultaneously testing a
user-defined selection of diversity indices. Molecular ecology resources 12, 1068–??78.
Evaluation of the methods in mcpHi11
Westfall, P. H. and Young, S. S. (1993) Resampling-Based Multiple Testing: Examples and Methods
Corresponding method sbdiv with method WYht
Besag, J., Green, P. J., Higdon, D., Mengersen, K. (1995) Bayesian computation and stochastic
Corresponding method sbdiv with method rpht
Corresponding method sbdiv with method tsht
Fritsch, K. S., Hsu, J. C. (1999) Multiple comparison of entropies with application to dinosaur
617–625.
Corresponding method sbdiv with method asht
Jost, L. (2008) G(ST) and its relatives do not measure differentiation. Molecular Ecology, 17,
4015-4026.
Corresponding method mcpi11

| asht | Internal function for simultaneous asymptotic intervals |

Description
Internal function for simultaneous asymptotic intervals

Note
Only internal function. Use function sbdiv instead
References


---

**Bacteria**

**Relative Abundances of Soil Bacteria**

**Description**

Relative abundances of soil bacteria from 27 samples collected in nine forest and 18 grassland sites in Germany. The data set includes abundances of 18 bacterial phyla (including three candidate phyla) and five proteobacterial classes.

**Usage**

data(Bacteria)

**Format**

A data frame with 27 observations on the following 24 variables.

- Land use type: a factor with levels forest grassland
- Acidobacteria: a numeric vector
- Actinobacteria: a numeric vector
- Bacteroidetes: a numeric vector
- Chloroflexi: a numeric vector
- Cyanobacteria: a numeric vector
- Deinococcus-Thermus: a numeric vector
- Fibrobacteres: a numeric vector
- Firmicutes: a numeric vector
- Fusobacteria: a numeric vector
- Gemmatimonadetes: a numeric vector
- Nitrospira: a numeric vector
- OP11: a numeric vector
- Planctomycetes: a numeric vector
- Spirochaetes: a numeric vector
- Tenericutes: a numeric vector
- TM7: a numeric vector
- Verrucomicrobia: a numeric vector
Details

Relative abundances of 18 bacterial phyla (including three candidate phyla) and five proteobacterial classes (alpha, beta, gamma, delta and epsilon) from two ecological metagenomics studies (Will et al. 2010, Nacke et al. 2011). There are 27 observations altogether, nine of which stem from forest and 18 from grassland plots in Germany.

One goal of these investigations was to unravel differences in bacterial diversity and community composition between the land use types forest and grassland.

The bacteria’s relative abundances were determined by analyzing the V2-V3 region of the 16S rRNA gene via pyrosequencing-based DNA techniques.

Source


Examples

data(Bacteria)

str(Bacteria)

### Assess whether there is a difference in biodiversity and community composition species richness (Shannon index, Simpson index) between grassland and forest.
### Bootstrap times set to 50 due to example time settings

library(simboot)
mcpHill(dataf=Bacteria[,2:24], fact=Bacteria[,1], boots=50, qval=c(0,1,2))
**Boutrp**

*Internal function*

**Description**

Internal function for method `rpht` in function `sbdiv`

**Note**

Only for internal use.

---

**CCdrp**

*Internal function*

**Description**

Internal function for method `rpht` in `sbdiv`

---

**contrMat**

*Contrast Matrices*

**Description**

Computes contrast matrices for several multiple comparison procedures.

**Usage**

```r
ccontrMat(n, type = c("Dunnett", "Tukey", "Sequen", "AVE", "Changepoint", "Williams", "Marcus", "McDermott", "UmbrellaWilliams", "GrandMean"), base = 1)
```

**Arguments**

- `n` a (possibly named) vector of sample sizes for each group.
- `type` type of contrast.
- `base` an integer specifying which group is considered the baseline group for Dunnett contrasts.

**Details**

Computes the requested matrix of contrasts for comparisons of mean levels.
Value

The matrix of contrasts with appropriate row names is returned.

Note

Function contrMat is adapted from package multcomp

References


Examples

```r
n <- c(10, 20, 30, 40)
names(n) <- paste("group", 1:4, sep="")
contrMat(n) # Dunnett is default
contrMat(n, base = 2) # use second level as baseline
contrMat(n, type = "Tukey")
contrMat(n, type = "Sequen")
contrMat(n, type = "AVE")
contrMat(n, type = "Changepoint")
contrMat(n, type = "Williams")
contrMat(n, type = "Marcus")
contrMat(n, type = "McDermott")
## Umbrella-protected Williams contrasts, i.e. a sequence of Williams-type contrasts with groups of higher order
## stepwise omitted
contrMat(n, type = "UmbrellaWilliams")
## comparison of each group with grand mean of all groups
contrMat(n, type = "GrandMean")
```

**corrmatgen**

*Internal function.*

Description

Correlation matrix for confidence intervals assuming multivariate standard normal distribution. Calculates the correlation matrix for method asci in function sbdiv

Usage

corrmatgen(CM, varp)

Arguments

- **CM** a matrix of contrast coefficients, dimension MxI, where M=number of contrasts, and I=number of groups in a oneway layout
- **varp** a numeric vector of groupwise variance estimates (length = I)
estShannonf

**Description**
Estimation function for Shannon’s index. Internal use in estShannonf.

**Usage**
estShannonf(x)

**Arguments**
x Vector of discrete-scaled numerical values.

**Details**
Estimator of Shannon-Wiener index with bias correction. Number of Species S in the bias correction does not take zeros into account.

**Value**
Shannon-Wiener index with bias correction

---
estShannonf

**Description**
Estimation function for Shannon’s index. Internal use in sdiv for methods rpht, tsht, asht. Sums up species counts in each columns for every treatment group and estimates Shannon’s index with bias correction on the resulting vectors of summed up species counts.

\[ \hat{H}_{BC_i} = \hat{H}_i + (S_i - 1)/(2N_{i\bullet}) - (1 - \sum (1/\hat{p}_{i\bullet s}))/(12N_{i\bullet}^2) - \sum ((1/\hat{p}_{i\bullet s}) - (1/\hat{p}_{i\bullet s}^2))/(12N_{i\bullet}^3); \]

\[ i = 1, ..., k; s = 1, ..., S; \hat{p}_{i\bullet s} = \frac{\sum_{n=1}^{s} x_{js}}{N_{i\bullet}}; \]

\[ \hat{H}_i = (-1) \sum_{s=1}^{S} (\hat{p}_{i\bullet s} \log(\hat{p}_{i\bullet s})) \]

\[ N_{i\bullet} = \sum_{j=1}^{n} N_{ij} \text{ Number of observed individuals in treatment } i. \]
estShannonWY

Usage

estShannonWY(x, f)

Arguments

x
n times p matrix containing species in p columns and replicates in n rows.
f
Factor variable containing treatment groups. Must be of length: replicates times treatment groups.

Value

estimate
Estimated Shannon-Wiener index for treatment groups
varest
Estimated variance of Shannon-Wiener index for treatment groups

estShannonWY
Estimator for Shannon’s index row wise.

Description

Estimation function for Shannon’s index. Internal use in wyht. Calculates Shannon-Wiener index with bias correction

$$\hat{H}BC_{ij} = \hat{H}_{ij} + (S_j - 1)/(2N_j) - (1 - \sum_{s=1}^{S_j}(1/\hat{p}_{ij,s}))/ (12N_j^2) - \sum_{s=1}^{S_j}((1/\hat{p}_{ij,s}) - (1/\hat{p}_{ij,s}^2))/ (12N_j^3);$$

$$\hat{H}_{ij} = (-1) \sum_{s=1}^{S_j}(\hat{p}_{ij,s}log(\hat{p}_{ij,s}))$$

i = 1, ..., k; j = 1, ..., n; s = 1, ..., S;
S_j = Number of observed species in replicate j;
N_j = Number of observed individuals in replicate j
for every row in a n \times p matrix.

Usage

estShannonWY(x)

Arguments

x
Vector of p numerical species counts.

Value

Shannon-Wiener index with bias correction
estSimpson  

Estimator for Simpson’s index

Description

Estimation function for Simpson’s index $1 - p^2 \times n/(n - 1)$. Internal use in estSimpson.

Usage

estSimpson(x)

Arguments

x  
Vector of discrete-scaled numerical values.

Value

Estimator of Simpson’s index

estSimpsonf  

Estimator for Simpson’s index ordered by a factorial variable f.

Description

Estimation function for Simpson’s index. Internal use in sbdiv for methods rpht, tsht, asht. Sums up species counts in each columns for every treatment group and estimates Simpson’s index on the resulting vectors of summed up species counts.

Usage

estSimpsonf(X, f)

Arguments

X  
n times p matrix containing species in p columns and replicates in nrows.

f  
Factor variable containing treatment groups. Must be of length: replicates times treatment groups.

Value

estimate  
Estimated Simpson index for treatment groups

varest  
Estimated variance of Simpson’s index for treatment groups
**estThetaRow**

*Internal function*

**Description**

Internal function for method *Wyht* in function *sbdiv*. Calculates the specified diversity index for every replicated sample in each treatment group.

**Usage**

```
estThetaRow(X, f, theta)
```

**Arguments**

- `X`: Matrix with dimension $n \times p$.
- `f`: Factorial variable containing treatment groups.
- `theta`

---

**mcpHill**

*Multiplicity-adjusted p-values for comparing biodiversity via simultaneous inference of a user-defined selection of diversity indices*

**Description**

The function `mcpHill` allows for comparing biological diversities of two or more groups. It simultaneously tests a user-defined selection of Hill numbers of orders $q$, which are considered appropriate and useful indices for measuring diversity (Jost 2008). As an output `mcpHill` gives p-values adjusted for multiplicity according to the method of Westfall & Young (1993).

**Usage**

```
mcpHill(dataf, fact, align = FALSE, block, boots = 5000, umat = FALSE, usermat, matty = "Dunnett", dunbase = 1, qval = seq(-1, 3), opt = "two.sided")
```

**Arguments**

- `dataf`: Data frame containing numerical values (e.g. species counts or relative abundances). Rows represent repeated observations of the (two or more) groups, columns represent taxonomic units (usually species, or phyla, classes etc.).
- `fact`: Vector assigning (two or more) factor levels to the observations, i.e. the groups to be compared. The length of `fact` must equal the number of rows in `dataf`.
- `align`: Logical indicating whether a block alignment should be carried out. If `TRUE`, the blocks must be specified as a vector in `block`. Default is `FALSE`. 
block Vector assigning which block an observation belongs to. Only required if align=TRUE. The length of block must equal the number of rows in dataf.

boots Number of bootstrap replications. Values lower than 999 are rejected. Default is 5000.

udmat Logical indicating whether user-defined contrasts are applied for multiple testing. If TRUE, a contrast matrix has to be specified via usermat. Default is FALSE, meaning that the contrast matrix is specified by a catchword (e.g. "Tukey", "Dunnett" etc.).

usermat Matrix specifying user-defined multiple testing contrasts. Only required if udmat=TRUE. The row sums in the matrix must equal zero.

mattype Type of contrast matrix for multiple comparisons of groups. Hence only required for comparisons of more than two groups. Can be specified by the catchwords used in function `contrMat` (e.g. "Dunnett", "Tukey", "GrandMean", "AVE", "Williams", "Changepoint" etc.). Default is "Dunnett".

dunbase Integer determining the factor group (in alphanumerical order) to be considered the baseline or control and therefore only needed for Dunnett-type multiple contrasts. Default is 1.

qval Vector containing the requested selection of q-values in order to specify the Hill numbers of orders q to be investigated. Default is `seq(-1,3)`.

opt "greater" performs an upper-tailed test, "less" a lower-tailed test and "two.sided" a two-tailed test. Default is "two.sided".

Value
The output of mcpHill is a matrix containing the chosen selection of Hill numbers (their orders q) in the first column. The multiplicity-adjusted p-values for each hypothesis tested are in the second column. The names of the rows denote which groups are being compared.

Author(s)
Philip Pallmann

References


Examples
```r
### Multiple testing with user-defined contrasts after block alignment
data(predatGM)
```
predatGM

Description

In a field trial with 8 complete blocks, one genetically modified crop variety and three varieties without genetic modification (S1, S2, S3) have been cultivated. Note that S1 is genetically closely related to the GM variety, and mainly differs from GM by not containing the transformation, while S2 and S3 are conventional varieties, which are genetically not closely related to GM and S1. In each of the 24 plots, a certain taxonomic group of predatory insects has been trapped. Trapped individuals have been classified to the species level. A total of 33 different species has been observed. For each plot, the summed counts of each species over one cultivation period is given in the variables Sp1, Sp2,...,Sp33. Among others, one question in research was: Does the genetic modified variety effect biodiversity of the (ecologically important, non-target) species?

Usage

data(predatGM)

Format

A data frame with 32 observations on the following 35 variables.

Block a numeric vector, values 1,...,8 indicate the blocks of the trial
Variety a factor distinguishing the four varieties in the field trial, with levels GM (the genetically modified variety), S1 (the near-isogenic, conventional variety), S2 and S3 (further conventional varieties)
Sp1 a numeric vector, observed counts of species 1
Sp2 a numeric vector, ...
Sp3 a numeric vector
Sp4 a numeric vector
Sp5 a numeric vector

mymat <- rbind("GM - S1" = c(1,-1,0,0), "GM - S2" = c(1,0,-1,0), "GM - S3" = c(1,0,0,-1), "S1 - S2" = c(0,1,-1,0), "S1 - S3" = c(0,1,0,-1) )

# example runs with only 100 bootstrap steps. For estimation use 2000 or more.
mcpHill(data=predatGM[,3:35], fact=predatGM[,2], align=TRUE, block=predatGM[,1], boots=100, usermat=mymat, qval=seq(-1, 3, by=0.5))

# with Dunnett-type contrast matrix
mcpHill(data=predatGM[,3:35], fact=predatGM[,2], align=TRUE, block=predatGM[,1], boots=100, usermat=FALSE, mattype = "Dunnett", qval=seq(-1, 3, by=0.5))
Sp6 a numeric vector
Sp7 a numeric vector
Sp8 a numeric vector
Sp9 a numeric vector
Sp10 a numeric vector
Sp11 a numeric vector
Sp12 a numeric vector
Sp13 a numeric vector
Sp14 a numeric vector
Sp15 a numeric vector
Sp16 a numeric vector
Sp17 a numeric vector
Sp18 a numeric vector
Sp19 a numeric vector
Sp20 a numeric vector
Sp21 a numeric vector
Sp22 a numeric vector
Sp23 a numeric vector
Sp24 a numeric vector
Sp25 a numeric vector
Sp26 a numeric vector
Sp27 a numeric vector
Sp28 a numeric vector
Sp29 a numeric vector
Sp30 a numeric vector
Sp31 a numeric vector
Sp32 a numeric vector
Sp33 a numeric vector

Source

Data set provided by Kai U. Priesnitz, Bavarian State Research Center for Agriculture, Institute for Plant Protection, Freising, Germany.
Examples

data(predatGM)

str(predatGM)

# Display data as a mosaicplot

# load("D:/Mueller/Biodiv/data/predatGM.rda")

# Matrix of counts with appropriate names
COUNTS<-as.matrix(predatGM[,3:35])
SPECNAM<-names(predatGM)[3:35]  
colnames(COUNTS)<-SPECNAM  
rownames(COUNTS)<-predatGM[,"Variety"]

# Assign colors and order by decreasing total abundance
COLX<grey(c(0,2,4,6,8,1,3,5,7)/8)
DMO<-COUNTS[,order(colSums(COUNTS), decreasing=TRUE)]
rownames(DMO)[15:33]<-"

# Mosaicplot
par(mar=c(4,2,1,1))
mosaicplot(DMO, col=COL, las=2, off=15, main="", cex=1.1)
mtext("A", side=3, line=-1.5, adj=0, cex=2)

rph

Internal function for simultaneous bayesian bootstrap intervals

Description

Internal function for simultaneous bayesian bootstrap intervals

Note

Only internal function. Use function sbdiv instead

References

Abundance data of Diptera with saprophagous larvae

Description

In a field trial with 6 complete blocks, three treatments have been applied: a genetically modified crop variety was cultivated without insecticide treatment (GM), its near-isogenic counterpart (i.e. not genetically modified but otherwise genetically closely related to the GM crop) has been cultivated without insecticide treatment (Iso), and the near-isogenic variety has been cultivated with insecticide treatment (Ins). In each of the 18 plots, two emergence traps have been placed and Diptera with saprophagous larvae were classified to the species level and counted. A total number of 25 different species has been observed and included in the present data set. For each plot, the summed counts of each species over one cultivation period (in 2002) and the two traps is given in the columns Acor, ..., Tnud. Among others, one question in this trial was: Does the genetic modified variety effect biodiversity of the (ecologically important, non-target) species in comparison to the isogenic variety (as a negative control) and in comparison to the insecticide treated plants (as a positive control)?

Usage

data(saproDipGM)

Format

A data frame with 18 observations on the following 27 variables.

- **Block** a numeric vector, values 1,...,6 indicate the blocks of the trial
- **Variety** a factor, distinguishing the 3 treatment levels: GM (genetically modified, no insecticide), Ins (not genetically modified, insecticide treatment), and Iso (not genetically modified, no insecticide)
- **Acor** a numeric vector of counts of the first species
- **Arub** a numeric vector
- **Aaph** a numeric vector
- **Bbre** a numeric vector
- **Btri** a numeric vector
- **Burt** a numeric vector
- **Bvag** a numeric vector
- **Bill** a numeric vector
- **Ccru** a numeric vector
- **Cmir** a numeric vector
- **Cvag** a numeric vector
- **Dnit** a numeric vector
- **Dand** a numeric vector
**SaprodipGM**

Lcin a numeric vector
Lcas a numeric vector
Malt a numeric vector
Moli a numeric vector
Mluc a numeric vector
Mt ox a numeric vector
Ppha a numeric vector
Sato a numeric vector
Spal a numeric vector
Sate a numeric vector
Sleu a numeric vector
Tnud a numeric vector

**Source**

Data set provided by Dr. Sabine Prescher, Institute for Biosafety of Genetically Modified Plants, Julius-Kuehn-Institut, Braunschweig, Germany

**Examples**

data(saprodipGM)
str(saprodipGM)

# load("D:/Mueller/Biodiv/data/saprodipGM.rda")

# Display data as a mosaicplot

# Matrix of counts with appropriate names
COUNTS<-as.matrix(saprodipGM[,3:27])
SPE CNAM<-names(saprodipGM)[3:27]
colnames(COUNTS)<-SPE CNAM
rownames(COUNTS)<-saprodipGM[,"Variety"]

# Assign colors and order by decreasing total abundance
COL<-grey(c(0,2,4,6,8,1,3,5,7)/8)
DMO<-COUNTS[,order(colSums(COUNTS), decreasing=TRUE)]

# Mosaicplot
par(mar=c(4,2,1,1))
mosaicplot(DMO, col=COL, las=2, off=15, main="", cex=1.1)
mtext("A", side=3, line=-1.5, adj=0, cex=2)
**sbdiv**

Perform simultaneous confidence intervals or adjusted p–values for the Shannon and the Simpson index.

---

**Description**

Function `sbdiv` estimates simultaneous confidence intervals for the Shannon or the Simpson index. This function provides calculation of several pre–defined contrasts for confidence intervals. Further self-defined contrast are applicable. Simultaneous resampling confidence intervals are estimated according to the Algorithm of Besag et al. (1995) using method `rpht`, Westfall et al. (1993) using method `wyht` or similar to Beran (1988) using method `tsht`. Further estimation of simultaneous asymptotic intervals adjusting for heterogeneous variances is provided by method `asht` according to Fritsch and Hsu (1999) and Rogers and Hsu (2001). However, estimation of asymptotic intervals may make no sense in data sets with replicated samples due to overdispersion.

**Usage**

```r
sbdiv(X, f, theta = c("Shannon", "Simpson"),
        type = c("Dunnett", "Tukey", "Sequen", "AVE",
                 "Changepoint", "Williams", "Marcus",
                 "McDermott", "UmbrellaWilliams", "GrandMean"),
        cmat = NULL, method = c("wyht", "tsht", "rpht", "asht"),
        conf.level = 0.95, alternative = c("two.sided", "less", "greater"),
        R = 2000, base = 1, ...)
```

**Arguments**

- **X**: Data frame containing numerical values for counts in columns. Every column represents on species.
- **f**: Vector of factorial variables for treatment groups. Vector length must be equal to the length of treatment groups multiplicated with sample replications.
- **theta**: Biodiversity index. Options are Shannon and Simpson index.
- **type**: Type of comparison. Options are Dunnett, Tukey, Sequen, AVE, Changepoint, Williams, Marcus, McDermott, UmbrellaWilliams, GrandMean intervals. We tested only Dunnett and Tukey contrasts in simulations.
- **cmat**: Optional self-defined contrast matrix. In case of using this argument, the type argument is not considered.
- **method**: Possible methods are simultaneous bootstrap confidence intervals: `wyht`, `tsht`, `rpht` and asymptotic simultaneous confidence intervals: `asht`. Adjusted and unadjusted p–values are estimated with method `wyht` and method `tsht`.
- **conf.level**: Pre-defined overall confidence level. Default is 0.95, while two-sided inference is estimated with $(1 - \text{conf.level})/2$ for each side and one-sided inference is estimated with $1 - \text{conf.level}$ for the side of interest.
- **alternative**: Specified type of interval. Could be "one-sided" or "two.sided".
Number of bootstrap steps. Default is 2000, which is a good compromise between accuracy and computing time

Control group. base = 1 uses the first group in alphabetical order.

Further optional arguments for the internal used function boot from package boot. Most importantly, the number of Bootstrap samples can be chosen via the parameter R (default is R=2000); see ?boot for further options.

Details

sbdiv is the main function for estimating the different multiplicity adjusted confidence intervals. Different methods are called from internal functions.

Value

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>conf.int</td>
<td>estimate: Estimated difference between groups. Estimators differ between the methods due to calculation. lower: Lower bounds of estimated intervals. upper: Upper bounds of estimated intervals.</td>
</tr>
<tr>
<td>p.value</td>
<td>adj. p: multiplicity adjusted p-values. raw p: unadjusted p-values</td>
</tr>
<tr>
<td>conf.level</td>
<td>Pre-specified confidence level</td>
</tr>
<tr>
<td>alternative</td>
<td>Pre-specified alternative</td>
</tr>
</tbody>
</table>

Author(s)

Ralph Scherer

References


Evaluation of the methods in sbdiv


Corresponding method sbdiv with method Wyht


Corresponding method sbdiv with method rpht


Corresponding method sbdiv with method tsht


Corresponding method sbdiv with method asht
## Examples

For plots of the datasets see the help files for the data sets.

### First dataset
```r
data(predatGM)

str(predatGM)

# remove block variable

datspec_1 <- predatGM[, -1]
str(datspec_1)

# Order of factorial variable

datspec_1$Variety
```

### Second dataset
```r
data(saprodipGM)

str(saprodipGM)

# remove block variable

datspec_2 <- saprodipGM[, -1]
str(datspec_2)

# Order of factor variable

datspec_2$Variety
```

## Notes

- ## Argument base = 1 uses GM as control group. Not directly executable
- ## Due to intensive computing time
  - sbdiv(X = datspec_1[, 2:length(datspec_1)], f = datspec_1[, 1], theta = "Shannon", type = "Dunnett", method = "WYht", conf.level = 0.95, alternative = "two.sided", R = 2000, base = 1)
  - Directly executable but senseless value for boot steps R

- ## Argument base = 2 uses Ins as control group. Not directly executable
- ## Due to intensive computing time
  - sbdiv(X = datspec_2[, 2:length(datspec_2)], f = datspec_2[, 1], theta = "Shannon", type = "Dunnett", method = "rpht", conf.level = 0.95, alternative = "two.sided", R = 2000, base = 2)
  - Directly executable but senseless value for boot steps R
Internal function

Description
Interval estimation in method rpci in function sbci

Note
Internal function. Use sbdiv instead.

---

Internal function for Simpson estimator

Description
Calculates Simpson’s index on probability vector \( p \)

Usage
Simpson(p)

Arguments
\( p \) Probability vector \( x_n/n \)

Value
Simpson’s index

Note
Only for internal use
Description

Internal function for simultaneous bootstrap intervals based on summed up counts for every species.

Note

Only internal function. Use function sbdiv instead.

References


Description

Internal function for Wald intervals in method asht in function sbdiv

Note

Internal function. Use function sbdiv instead.

References

Index

*Topic **datasets**
  Bacteria, 4
  predatGM, 13
  saproDipGM, 16

*Topic **htest, nonparametric, multivariate**
  simboot-package, 2

*Topic **htest**
  asht, 3
  mcpHill, 11
  rph, 15
  sbdiv, 18
  tsht, 22
  WYHt, 22

*Topic **misc**
  Boutrp, 6
  CCDrp, 6
  contrMat, 6
  cormatgen, 7
  estShannon, 8
  estShannonf, 8
  estShannonWY, 9
  estSimpson, 10
  estSimpsonf, 10
  estThetaRow, 11
  SCIrp, 21
  Simpson, 21
  waldci, 22

asht, 3, 8, 10, 18, 19, 22
Bacteria, 4
Boutrp, 6
CCDrp, 6
contrMat, 6, 12
cormatgen, 7

estShannon, 8
estShannonf, 8, 8
estShannonWY, 9
estSimpson, 10
estSimpsonf, 10
estThetaRow, 11
mcPHill, 3, 11
predatGM, 13
rph, 3, 6, 8, 10, 15, 18, 19
saproDipGM, 16
sbdiv, 3, 6–8, 10, 11, 15, 18, 19, 21, 22
SCIrp, 21
simboot (simboot-package), 2
simboot-package, 2
Simpson, 21
tsht, 3, 8, 10, 18, 19, 22
waldci, 22
WYHt, 3, 9, 11, 18, 19, 22

23