

# Package ‘RInSp’

January 24, 2018

**Type** Package

**Title** R Individual Specialization (RInSp)

**Version** 1.2.3

**Date** 2018-01-24

**Author** Dr. Nicola Zaccarelli, Dr. Giorgio Mancinelli, and Prof. Dan Bolnick

**Maintainer** Dr. Nicola Zaccarelli <nicola.zaccarelli@gmail.com>

**Description** Functions to calculate several ecological indices of individual and population niche width (Araujo's E, clustering and pairwise similarity among individuals, IS, Petraitis' W, and Rough-garden's WIC/TNW) to assess individual specialization based on data of resource use. Resource use can be quantified by counts of categories, measures of mass/length or proportions. Monte Carlo resampling procedures are available for hypothesis testing against multinomial null models.

**License** GPL (>= 2)

**NeedsCompilation** yes

**Repository** CRAN

**Encoding** UTF-8

**LazyData** true

**Date/Publication** 2018-01-24 16:38:52 UTC

## R topics documented:

RInSp-package	2
Eindex	4
Emc	7
Hier2L	10
import.RInSp	11
Likelihood and Wi	13
MCp.RInSp	15
NODF	16
Null.Hp.RInSp	18
overlap	20

pop.diet . . . . .	21
PSicalc . . . . .	23
RInSp2vegan . . . . .	25
Stickleback . . . . .	26
sumMC.RInSp . . . . .	27
Trout . . . . .	28
WTcMC . . . . .	29
WTdMC . . . . .	31
<b>Index</b>	<b>34</b>

---

RInSp-package	<i>R Individual Specialization</i>
---------------	------------------------------------

---

## Description

“R Individual Specialization” (RInSp) is an R package for analysing resource use data to calculate ecological measures of within-population niche variation based on data on individuals’ resource use (e.g., stomach contents, repeated observations of prey choice or habitat use). The package implements several ecological niche metrics to measure individual and population niche width in order to measure individual specialization. This version of RInSp is able to compute Araujo’s E, measures of clustering and pairwise similarity among individuals, IS, Petraitis’ W, and Roughgarden’s WIC/TNW, and an example of a measure of nestedness (NODF). For some of these indices Monte Carlo re-sampling procedures for testing significance are provided.

To cite package check the command CITATION or use:

Zaccarelli, N., Mancinelli, G. and Bolnick, D.I. (2013) RInSp: an R package for the analysis of individual specialisation in resource use. *Methods in Ecology and Evolution*, 4(11): 1018-1023, doi: 10.1111/2041-210X.12079

This package can be used on Mac, Windows and UNIX platforms, and on both i386 and x64 architectures. It is based on IndSpec1.0 developed by Prof. D. Bolnick (Bolnick et al. 2002), and on Dieta1 developed by Dr. M.A. Araujo (Araujo et al. 2008).

The original source code of IndSpec1.0 is available from the “Ecological Archives” of the Ecological Society of America (<http://esapubs.org/archive/>): identifier E083-056-S1; <http://esapubs.org/archive/ecol/E083/056/>. The original source code Dieta1 is available from the “Ecological Archives” of the Ecological Society of America (<http://esapubs.org/archive/>): identifier E089-115-A1; <http://esapubs.org/archive/ecol/E089/115/>.

Version 1.1 of the package fixes three issues:

1. in `Eindex` the case of highly specialised individuals is now handled correctly;
2. in `like.Wi` the probability for each value of Petraitis’ W is properly handled;
3. in `Emc` the C code taken from Dieta1.c has been fixed for an error in the calculation of the weight matrix when highly specialised individuals are present.

Version 1.2.3 of the package has been changed to address the following things:

1. we introduce a filter to remove from the PS matrix all values which are lower than a set threshold (i.e., the parameter "*precision*"). In this way some interactions are removed due to their negligible weight in the PS matrix. The affected functions are: `Eindex`, `Emc`, `PSicalc` and the C code used for the resampling;
2. in `Eindex` users are now warned for cases of isolated individuals by means of a text message and by the value of the variable `Isolation`;
3. all progressbars are removed from all functions, so in case your problem is big have a little bit of patience;
4. now `Eindex` and `Emc` will work correctly with the Barrat index in case of isolated individuals;
5. Undefined global functions or variables are now properly registered in the NAMESPACE file;
6. native routines have been registered and symbol search has been disabled;
7. unprotected variables have been taken care of.

The Authors thank for their help in improving the package Miss Georgia Karoline Kosmala (<geh\_karol@hotmail.com>), Miss Ambika Kamath (<geh\_karol@hotmail.com>), Miss Isabel Fisk Baruque (<isabel.fiskbaruque@gmail.com>) and Dr. Marcio S. Araujo (<msaraujo@rc.unesp.br>).

## Details

Package: RInSp  
Type: Package  
Version: 1.2.3  
Date: 2018-24-01  
License: GPL (>= 2)

## Author(s)

Dr. Nicola ZACCARELLI <nicola.zaccarelli@gmail.com>,  
Dr. Giorgio MANCINELLI <giorgio.mancinelli@unisalento.it>,  
and Prof. Dan BOLNICK <danbolnick@mail.utexas.edu>

## References

- Almeida-Neto M., Guimaraes P., Guimaraes P.R., Loyola R.D., and Ulrich W. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**: 1227-1239.
- Araujo M.S., Guimaraes Jr., P.R., Svanback, R., Pinheiro, A., Guimaraes P., dos Reis, S.F., and Bolnick, D.I. 2008. Network analysis reveals contrasting effects of intraspecific competition on individual vs. population diets. *Ecology* **89**: 1981-1993.
- Araujo M.A., Layman C., and Bolnick D.I. 2011. The ecological causes of individual specialization. *Ecology Letters* **14**: 948-958.
- Aversen. J.N. 1969. Jackknifing U-statistics. *Annals of Mathematical Statistics* **40**: 2076-2100.

- Barrat A., Barthelemy M., Pastor-Satorras R., and Vespignani A. 2004. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences* **101**: 3747-3752
- Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.
- Bolnick, D.I., and Paull, J.S. 2009. Morphological and dietary differences between individuals are weakly but positively correlated within a population of threespine Stickleback. *Evolutionary Ecology Research* **11**: 1217-1233.
- Onnela J.P., Saramaki J., Kertesz J., and Kaski K. 2005. Intensity and coherence of motifs in weighted complex networks. *Physics Review E* **71**: 065103.
- Roughgarden, J. 1974. Niche width: biogeographic patterns among *Anolis lizard* populations. *American Naturalist* **108**: 429-411.
- Saramaki J., Kivela M., Onnela J.P., Kaski K., and Kertesz J. 2007. Generalizations of the clustering coefficient to weighted complex networks. *Physics Review E* **75**: 027105.
- Schoener, T.W. 1968. The *Anolis* lizards of Bimini: resources partitioning in a complex fauna. *Ecology* **49**: 704-726.
- Zaccarelli, N., Bolnick, D.I., and Mancinelli, G. 2013. RInSp: an R package for the analysis of individual specialization in resource use. *Methods in Ecology and Evolution* **4**: 1018-1023.

---

Eindex

---

*Calculate index E and Cws*


---

### Description

The procedure calculates the E measure of interindividual variation, its variance and the value of the  $C_{ws}$  measure of modularity after Araujo et al. (2008).

### Usage

```
Eindex(dataset, index = "saramaki", jackknife = FALSE, precision = 1e-9)
```

### Arguments

dataset	Object of class RInSp with data of type “double”, “integer” or “proportions”.
index	The type of clustering coefficient to use. Valid values are Saramaki’s and Barrat’s index: “saramaki” or “barrat”.
jackknife	Specify if a jackknife estimate of the index variance is required. Default is FALSE.
precision	a filter to remove from the PS matrix all values which are lower than "precision". In this way some interactions are removed due to their small weight in the PS matrix.

## Details

The index E has been proposed by Araujo et al. (2008) as a measure of individual specialization where, in absence of interindividual niche variation, its value is zero. The index will increase towards one with the increase of interindividual variation.

A jackknife estimation of the variance of E can be derived using the formalism of U-statistics (Aversen, 1969). For a complete description and a formal demonstration of the jackknife estimation of the variance of the E index the reader is referred to Araujo et al. (2008).

A measure of the relative degree of clustering in a network to test for modularity in the niche overlap network is  $C_{ws}$ . In a totally random network (i.e., a network consisting of individuals that sample randomly from the population niche),  $C_{ws}$  is approximately 0, indicating no modularity. If individuals form discrete groups specialized on distinct sets of resources,  $C_{ws} > 0$ , and the network is modular. If  $C_{ws} < 0$  the network degree of clustering is actually lower than what would be expected solely on the overall network density of connections, indicating that diet variation takes place at the level of the individual, as opposed to discrete groups.

The relative degree of clustering  $C_{ws}$  is obtained as the mean values over all nodes the niche overlap network of the individual node weighted clustering coefficients ( $C_{w_i}$ ).

The clustering can be measured using two different type of weighted clustering coefficients. In general the degree of unweighted clustering around a vertex  $i$  in a network is quantified by evaluating the number of triangles in which the vertex participates normalized by the maximum possible number of such triangles. Hence we have zero if none of the neighbours of a vertex are connected, otherwise we have one. By extending the above line of reasoning, the weighted clustering coefficient should also take into account how much weight is present in the neighbourhood of the vertex, compared to some limiting case. This can be done in several ways. Barrat et al. (2004) were the first to propose a weighted version of the clustering coefficient of the form:

$$C_{w_i} = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij}a_{ih}a_{jh}$$

where  $s_i$  is the sum of the weights ( $w_i$ ) of all the edges between node  $i$  and the nodes to which it is connected;  $k_i$  is the number of edges between node  $i$  and its neighbours;  $w_{ij}$  is the weight of the edge between the two nodes  $i$  and  $j$ ;  $a$  is 1 if an edge is present between each pair  $ij$ ,  $ih$ , and  $jh$  respectively, and zero otherwise. The summation, therefore, quantifies the weights of all edges between node  $i$  and its neighbours that are also neighbours to each other.

Samaraki et al. (2007) and Onnela et al. (2005) proposed a version of a clustering index of the form:

$$C_{w_i} = \frac{1}{k_i(k_i - 1)} \sum_{j,h} (w_{ij}w_{ih}w_{jh})^{\frac{1}{3}}$$

where  $k_i$  is the number of edges between individual  $i$  and its neighbours;  $w_{ij}$  is the weight of the edge between individual  $i$  and  $j$  obtained by dividing the actual weight by the maximum of all weights. The summation, therefore, quantifies the weights of all edges between individual  $i$  and its neighbours that are also neighbours to each other.

The default value of the procedure has been set for analogy with the Dieta1.c code provided by Araujo et al. (2008). The original source code Dieta1.c is available from the ‘‘Ecological Archives’’

of the Ecological Society of America (<http://esapubs.org/archive/>): identifier E089-115-A1; <http://esapubs.org/archive/eco1/E089/115/>.

Version 1.1 of the package fixes the case of highly specialised individuals in the calculation of  $C_{ws}$ .

### Value

The result is a list of class ‘RInSp’ composed of:

Omean	Mean value of the measure of the network overall degree of pairwise overlap.
E	Value of the index of individual specialization E.
PS	Matrix of the measure of niche pairwise overlap between i and j, adapted from Schoener (1968).
PSbinary	Binary matrix derived by applying the threshold of Omean to the PS matrix. This matrix can be imported into the software PAJEK ( <a href="http://vlado.fmf.uni-lj.si/pub/networks/pajek/">http://vlado.fmf.uni-lj.si/pub/networks/pajek/</a> ) to draw binary networks of diet similarity among individuals (e.g., see Araujo et al. 2008).
Ejack	Values of the measure of interindividual variation used for the Jackknife estimates of Var(E).
VarE	Jackknife estimates of the variance of E.
CW	is the network weighted clustering coefficient.
CwS	Vector of the individuals weighted clustering coefficients.
Cw	Value of the measure of modularity.
index	The type of clustering coefficient used.
Ki	A vector with the degree of the nodes of the network.
Precision	The value of the selected threshold for precision.
Isolation	If the value is different from zero there are isolated individuals in the network.

### Author(s)

Dr. Nicola ZACCARELLI

### References

- Araujo M.S., Guimaraes Jr., P.R., Svanback, R., Pinheiro, A., Guimaraes P., dos Reis, S.F., and Bolnick, D.I. 2008. Network analysis reveals contrasting effects of intraspecific competition on individual vs. population diets. *Ecology* **89**: 1981-1993.
- Aversen J.N. 1969. Jackknifing U-statistics. *Annals of Mathematical Statistics* **40**: 2076-2100.
- Barrat A., Barthelemy M., Pastor-Satorras R., and Vespignani, A. 2004. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences* **101**: 3747-3752
- Onnela J.P., Saramaki J., Kertesz J., and Kaski, K. 2005. Intensity and coherence of motifs in weighted complex networks. *Physics Review E* **71**: 065103.
- Saramaki J., Kivela M., Onnela J.P., Kaski K., and Kertesz, J. 2007. Generalizations of the clustering coefficient to weighted complex networks. *Physics Review E* **75**: 027105.
- Schoener, T.W. 1968. The *Anolis* lizards of Bimini: resources partitioning in a complex fauna. *Ecology* **49**: 704-726.

**See Also**

Function [Emc](#).

**Examples**

```
# Eindex example with data from Bolnick and Paull (2009)
data(Stickleback)
# Select a single spatial sampling site (site B)
GutContents_SiteB <- import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "B"))
# Index calculation with jackknife variance estimate
# This can take time for big data sets
Eresult <- Eindex(GutContents_SiteB, index = "saramaki", jackknife = TRUE)
rm(list=ls(all=TRUE))
```

---

Emc

*Monte Carlo resampling procedure for index E and Cws*


---

**Description**

The procedure performs a Monte Carlo resampling under a null hypothesis to calculate the E measure of interindividual variation and the value of the  $C_{ws}$  measure of modularity following Araujo et al. (2008).

**Usage**

```
Emc(dataset, popd.type = "sum", index = "saramaki", replicates=999, precision = 1e-9)
```

**Arguments**

dataset	Object of class RInSp. Only data of type “integer” can be used. See <a href="#">import.RInSp</a> for details.
popd.type	Option specifying the way to calculate the population diet. Valid values are “sum” or “average”. See <a href="#">pop.diet</a> for details.
index	The type of clustering coefficient to use. Valid values are Saramaki’s and Barrat’s index: “saramaki” or “barrat”. See <a href="#">Eindex</a> for details.
replicates	Number of replicates for the Monte Carlo resampling, under the null hypothesis that any observed diet variation arose from individuals sampling stochastically from a shared distribution.
precision	a filter to remove from the PS matrix all values which are lower than “precision”. In this way some interactions are removed due to their small weight in the PS matrix.

## Details

For counts of individual prey items representing approximately independent prey-capture decisions (i.e., integer data type), it is possible to run a Monte Carlo resampling simulation to test the null hypothesis that any observed diet variation arose from individuals sampling stochastically from a shared distribution. See Araujo et al. (2011) for a discussion of the biological assumptions underlying this Monte Carlo resampling procedure. In the simulations, each individual is assigned a number of prey items equal to the number of items it was observed eating, and then prey items are randomly assigned to the individual's diet via multinomial sampling from the observed population resource distribution. Next, both  $E$  and  $C_{ws}$  are recalculated for the resulting simulated population. The program generates a specified number of such null datasets, and the observed test statistic is compared to the distribution observed under the null hypothesis.

Note that the Monte Carlo method should not be applied to non-integer diet data (e.g., total mass of each prey species, proportion of stomach volume per prey species).

There are two options for calculating the population's diet proportions ( $q_j$ , the proportion of the resource  $j$  in the population's diet): *sum* and *average*. When *sum* is specified, one sums up all resource counts within a category across all individuals to get the population's use, then determine the proportion of each resource category in the population's repertoire:

$$q_j = \frac{\sum_i n_{ij}}{\sum_{ji} n_{ij}}$$

The drawback of this approach is that individuals that eat large numbers of items, or larger total mass of items, will bias the population to look more like them. The *average* method (average proportion) circumvents this problem by first converting individual diets into proportions  $p_{ik}$ , then averaging these proportions for each resource  $k$ .

An adjusted value of  $E$  is calculated as:

$$E_{adj} = \frac{E_{obs} - E_{null}}{1 - E_{null}}$$

This rescales the value of  $E$  (i.e., the observed value  $E_{obs}$ ) to range from a minimum of 0 when the observed value is equal to the mean Null value (i.e.,  $E_{null}$ , up to a maximum of 1 when individual specialization is strongest). This allows one to compare datasets with different mean Null values.

Version 1.1 of the package fixes the case of highly specialised individuals in the calculation of  $C_{ws}$  by removing a bug in the C code taken from `Dieta1.c` affecting how the weight matrix and the number of links were calculated when highly specialised individuals are present.

## Value

The result is a list of class 'RInSp' composed of:

$E$	Value of the index of individual specialization $E$ .
<code>meannullE</code>	The mean value of the Monte Carlo resampling simulated data.
<code>Eadj</code>	The adjusted value of $E$ .
<code>p.value</code>	It is the proportion of Monte Carlo resampling simulated values greater than the observed $E$ .



montecarlo	It is a matrix where the first line collects the values of Omean, E, CW and Cws for the provided dataset, and the following lines hold the Monte Carlo resampled values.
pop.diet	Option specifying the way to calculate the population diet. Valid values are “sum” or “average”.
ind.type	The type of clustering coefficient to use. Valid values are Saramaki’s and Barrat’s index: “saramaki” or “barrat”.
parameter	an index used in <code>sumMC.RInSp</code> to point to the column of the matrix ‘montecarlo’ for plotting results. If zero no Monte Carlo resampling is performed.
Precision	The value of the selected threshold for precision.

### Author(s)

Dr. Nicola ZACCARELLI. Please note that the C code is partially based on the C source of Dieta.c by Araujo et al. (2008). The original source code Dieta1.c (identifier E089-115-A1; <http://esapubs.org/archive/ecol/E089/115/>) can be downloaded from the “Ecological Archives” of the Ecological Society of America (<http://esapubs.org/archive/>)

### References

- Araujo M.S., Guimaraes Jr., P.R., Svanback, R., Pinheiro, A., Guimaraes P., dos Reis, S.F., and Bolnick, D.I. 2008. Network analysis reveals contrasting effects of intraspecific competition on individual vs. population diets. *Ecology* **89**: 1981-1993.
- Araujo M.A., Layman C., and Bolnick, D.I. 2011. The ecological causes of individual specialization. *Ecology Letters* **14**: 948-958.
- Schoener, T.W. 1968. The *Anolis* lizards of Bimini: resources partitioning in a complex fauna. *Ecology* **49**: 704-726.

### See Also

Function `Eindex`, `Null.Hp.RInSp`, `sumMC.RInSp`.

### Examples

```
# Example with stickleback data from Bolnick and Paull 2009
data(Stickleback)
# Select a single spatial sampling site (site B)
GutContents_SiteB <- import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "B"))
# Warning, the number of replicates is set low to speed up the example!
# Note, for real analyses we recommend to set replicates => 999
ResEmc <- Emc(GutContents_SiteB, popd.type = "average", replicates = 249)

# This Eadj re-scales the observed E value to range from 0 when
# it is equal to the null expectation, up to a maximum of 1.
# This allows comparison of E values across samples with
# different null Es.
# For example, consider a case with no individual specialization
```

```

Nulldata <- Null.Hp.RInSp(GutContents_SiteB, prop="average")
# Warning, the number of replicates is set low to speed up he example!
# Note, for real analyses we recommend to set replicates => 999
EmcNull <- Emc(Nulldata, popd.type = "average", replicates = 249)
rm(list=ls(all=TRUE))

```

---

Hier2L

*Variance partition and WIC/TNW*


---

### Description

Variance partition based on a single multi-level factor and comparison of the relative contribution of the “*Between Group Component*” (BGC) versus the ratio WIC/TNW. For a more general discussion of variance partition the reader is pointed to Viola et al. (2012).

### Usage

```
Hier2L(dataset, factor = 1, weight.type = "N_items")
```

### Arguments

dataset	Object of class RInSp with data of type “double”.
factor	Numeric value indicating the column in the <i>info</i> part of the RInSp input dataset containing the factor used for build subsets.
weight.type	The weighting option for unbalanced prey numbers. See details in <a href="#">WTcMC</a> .

### Details

The Beteen Group Component is calculated as the (population) variance of the mean values of the n-levels groups.

The level of checking performed by the procedure is low. Pay attention to input parameters.

### Value

The result is a matrix of 5 rows by n+1 columns, when n is the number of levels of the factor used to create groups in the dataset. The first four rows are the values for WIC, BIC, TNC and WIC/TNW for all possible levels and the whole dataset. Last row is the value of the “*Between Group Component*” (BGC) for the n-levels factor.

### Author(s)

Dr. Nicola ZACCARELLI and Dr. Giorgio MANCINELLI

### References

Violle C., Enquist B.J., McGill B.J., Jiang L., Albert C.H., Hulshof C., Jung V. and Messier J. 2012. The return of the variance: intraspecific variability in community ecology. *Trends in Ecology & Evolution* **27**: 244252.

**See Also**

Function [WTcMC](#).

**Examples**

```
# Example of variance decomposition
data(Trout)
TroutRIS= import.RInSp(Trout, col.header=TRUE, row.names=1, info.cols=2, data.type="double")
decomp = Hier2L(TroutRIS, factor=1)
rm(list=ls(all=TRUE))
```

---

import.RInSp

---

*Import data into a RInSp object*


---

**Description**

The procedure reads and checks data to create an object of class RInSp.

**Usage**

```
import.RInSp(filename, col.header = FALSE, row.names = 0, info.cols = 0,
subset.column = 0, subset.rows = NA, data.type = "integer", print.messages=TRUE)
```

**Arguments**

filename	Name of the file or dataframe to be read.
col.header	Logical value to indicate if a header row is there. Default is FALSE.
row.names	Numeric value for column of rows' name. Default is zero for no names.
info.cols	A vector collecting columns numbers for additional information. Defaults is to have no additional information.
subset.column	A vector of columns' indices to be used as a subset. Default is to have no columns subsetting.
subset.rows	A string vector where the first element points to the column name to be used for rows subsetting, and following elements with criteria. Default is to have no rows subsetting.
data.type	Data type among "integer" or "double". From the data a proportion matrix will be produced. Default is to use integers/counts values.
print.messages	Prints messages concerning the number of rows and columns eventually deleted after subsetting because composed of all zeros. Default is TRUE.

## Details

Three different types of data can be used. Integer/count values and decimal/real values are stored in the “resources” section of the output list and used to derive all meaningful information. While for proportions the “resources” section will be empty. Valid key words are: “integer”, “double”, and “proportion”. Use zero for empty cells.

The procedure will check for the presence of zero sum columns/rows for the selected dataset. A warning is printed in case of column/row deletion. It must be considered that the degree of checking on the subsetting is low.

## Value

Return an list of class RInSp composed by:

resources	A matrix of the resources data.
proportions	A matrix of proportions of each resources in its row. This matrix can be imported into the software PAJEK ( <a href="http://vlado.fmf.uni-lj.si/pub/networks/pajek/">http://vlado.fmf.uni-lj.si/pub/networks/pajek/</a> ) to draw a weighted bipartite network connecting individuals to the various prey categories.
data.type	Data type used.
col.names	Name of the different columns of the resource data.
ind.names	Name of the individual.
info	A data frame containing additional information for the resource data.
num.prey	Number of resources/prey (i.e., columns) in the dataset after zero sum checking.
num.individuals	Number of individuals/sites (i.e., rows) in the dataset after zero sum checking.
num.zero.prey	Number of resources/prey (i.e., columns) in the dataset without zero sum checking.
num.ind.zero	Number of individuals/sites (i.e., rows) in the dataset without zero sum checking.

## Author(s)

Dr. Nicola ZACCARELLI

## Examples

```
data(Stickleback)
# Import data
GutContents = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13))
GutContents
# Select a single spatial sampling site (site A)
GutContents_SiteA = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13),
subset.rows = c("Site", "A"))

# Select a subset of prey types
GutContents_subset = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13),
```

```

subset.column = c(13:28, 45))

# Lump prey types into functional groups then import data
# define new columns representing lumped prey categories
attach(Stickleback)
Copepods <- Calanoid + Cyclopoid + Harpacticoid
Diptera <- Diptera.Pupae + Chironomid.larvae + Ceratopogonid.larvae + Tipulid.larvae +
Tipulidae.Adult + Diptera.Adult + Diptera.Larvae + Ceratopogonid.Adult
InsectLarvae <- Ephemeroptera + Trichoptera.larvae + Ephemeroptera.pupae +
Zygoptera.larvae + Plecoptera.larvae
Cladocera <- Bosmina + Polyphemus + Holopedium + Daphnia + Chydorus
names(Stickleback)
GutContents_lumped <- import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13),
subset.column = c(18,24,28,31,43:46))
rm(list=ls(all=TRUE))

```

---

Likelihood and Wi

*Likelihood measure of niche breadth*


---

### Description

The procedure calculates the estimation of likelihood measures of niche breadth and overlap described in Petraitis (1979).

### Usage

```
like.Wi(dataset)
```

### Arguments

dataset            Object of class RInSp with data.

### Details

The function returns the likelihood of the observed diet ( $\lambda_i$ ) the associated probability, and the value of the Petraitis' W. The likelihood of the observed diet of individual  $i$  is:

$$\lambda_i = \prod_j \left( \frac{q_j}{p_{ij}} \right)^{n_{ij}}$$

where  $q_j$  is the population proportion of the resource  $j$ ,  $p_{ij}$  is the proportion of the resource  $j$  in the diet of the individual  $i$ , and  $n_{ij}$  is the number of items for the individual  $i$  and the resource  $j$ .

This can be used to calculate a p-value to test the significance of the diet specialization, as  $-2\ln(\lambda)$  is distributed as a chi-square with  $(r-1)$  degrees of freedom, where  $r$  is the number of resource categories.

The generalised likelihood ratio test rejects the null hypothesis for a unilateral alternative hypothesis using significance level  $\alpha$  if:

$$-2\ln(\lambda) > \chi_{(r-1)}^2$$

Petraitis'  $W$  is computed following:

$$W_i = \lambda_i^{(1/D_i)}$$

where  $D_i$  is the number of diet items recorded in the diet of individual  $i$ . This index is a measure of niche width relative to a specified distribution. For a complete generalist individual,  $W_i = 1$ , and the value decreases with greater specialization.

### Value

Return a list of class `RInSp` with:

MeanWi	the mean population value of $W_i$ ;
ResCat	the number of resource categories;
ind.vals	A matrix with three columns: "Likelihood" with value of the likelihood index for the individual $i$ ; "p-value" for the the probability associated to the likelihood value; "Wi" with the value of the Petraitis' $W$ index for the individual $i$ .

### Author(s)

Dr. Nicola ZACCARELLI

### References

- Petraitis, P. S. 1979. Likelihood measures of niche breadth and overlap. *Ecology* **60**(4): 703-710.
- Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.

### Examples

```
# Likelihood and Wi example with stickleback data
# from Bolnick and Paull 2009
data(Stickleback)
# Select a single spatial sampling site (site D)
SiteD = import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "D"))
Wi = like.Wi(SiteD)
rm(list=ls(all=TRUE))
```

---

MCp.RInSp

*Monte Carlo resampling procedure*

---

### Description

The procedure replicates the two strategies of Monte Carlo resampling available in the following commands: `Emc`, `PSicalc`, `WTdMC`, and `WTcMC`.

### Usage

```
MCp.RInSp(dataset, pop.diet = "sum", replicates = 999)
```

### Arguments

<code>dataset</code>	Object of class <code>RInSp</code> with data of type “double” or “integer”.
<code>pop.diet</code>	Option specifying the type of calculation for the population diet: “sum” or “average”. See <code>pop.diet</code> for details.
<code>replicates</code>	Number of resampled data matrices.

### Details

See Araujo et al (2011) for a discussion of the biological assumptions underlying this Monte Carlo resampling procedure.

For counts of individual prey items representing approximately independent prey-capture decisions (i.e., integer data type), it is possible to run a Monte Carlo resampling simulation to test the null hypothesis that any observed diet variation arose from individuals sampling stochastically from a shared distribution. In the simulations, each individual is assigned a number of prey items equal to the number of items it was observed eating, and then prey items are randomly assigned to the individual’s diet via multinomial sampling from the observed population resource distribution.

For double data type, as mass of each food category in an individual’s gut, the Monte Carlo resampling procedure first determines the number of prey used by each individual,  $n_i$ . The program then randomly reassigns each individual  $n_i$  prey items drawn from the population distribution of items. This is repeated a user-specified number of times. Hence, the resampled population yields a null model corresponding to a population composed of generalists that sample randomly from the population’s diet, and have diet sizes equal to those of the observed data set.

The procedure can easily create a big three dimensional array. This object can create problems as it can consume an huge amount of memory. If needed use a package like “ff” and a for loop to create an array able to hold the results.

### Value

A matrix of “replicates + 1” datasets with the same number of individuals/resources of the input file. The first matrix is the original input data file.

### Author(s)

Dr. Nicola ZACCARELLI

## References

Araujo M.A., Layman C., and Bolnick D.I. 2011. The ecological causes of individual specialization. *Ecology Letters* **14**: 948-958.

## See Also

[pop.diet](#), [Emc](#), [PSicalc](#), [WTdMC](#), [WTcMC](#).

## Examples

```
# Monte Carlo resampling for NODF
# EXAMPLE NOT RUN
# with stickleback data from Bolnick and Paull 2009
# data(Stickleback)
# Select a single spatial sampling site (site A)
# SiteA = import.RInSp(Stickleback, row.names = 1,
# info.cols = c(2:13), subset.rows = c("Site", "A"))
# nreplicates = 999
# Null.data = MCp.RInSp(SiteA, replicates = nreplicates)
# Build a vector holding real and resampled results
# the NODF values and warning messages from import.RInSp
# are printed, too.
# null.d.NODF = c()
# for (i in 1:(nreplicates + 1)) {
#   null.d.NODF = c(null.d.NODF, NODF(import.RInSp(Null.data[ , , i], print.messages=FALSE),
#   print.results=FALSE)$NODF)
# }
# hist(null.d.NODF)
# summary(null.d.NODF)
# Observed value of NODF from data
# null.d.NODF[1]
# rm(list=ls(all=TRUE))
```

---

NODF

*Nestedness metric based on overlap and decreasing fill*

---

## Description

The procedure calculates the “nestedness metric based on overlap and decreasing fill” (NODF) for a binary matrix following Almeida-Neto et al. (2008).

## Usage

```
NODF(dataset, print.results= TRUE)
```

## Arguments

`dataset`            Object of class `RInSp` with data of type “double” or “integer”.  
`print.results`    Define if results for NODF should be printed. Default is `TRUE`.



## Details

Nestedness is a feature of binary matrices (also called presence/absence or incidence matrices). The notion of nestedness is particularly relevant for studies focusing the patterns of species occurrence among a set of locations (e.g., islands) and the patterns of interacting species within ecological networks.

Almeida-Neto et al. (2008) propose a nestedness metric is based on two simple properties: decreasing fill (or DF) and paired overlap (or PO). Assuming that in a matrix with  $m$  rows and  $n$  columns, row  $i$  is located at an upper position from row  $j$ , and column  $k$  is located at a left position from column  $l$ . In addition, let  $MT$  be the marginal total (i.e. the sum of 1's) of any column or row. For any pair of rows/columns  $i$  and  $j$ ,  $DF_{ij}$  will be equal to 100 if  $MT_j$  is lower than  $MT_i$ . Alternatively,  $DF_{ij}$  will be equal to 0 if  $MT_j$  is greater or equal to  $MT_i$ . For columns/rows, paired overlap ( $PO_{kl}$ ) is simply the percentage of 1's in a given column/row  $l$  that are located at identical row/column positions to those in a column/row  $k$ . For any left-to-right column pair and, similarly, for any up-to-down row pair, there is a degree of paired nestedness ( $N_{paired}$ ) as zero if  $DF_{paired}$  is zero, and PO if  $DF_{paired}$  is 100.

From the  $n(n-1)/2$  and  $m(m-1)/2$  paired degrees of nestedness for  $n$  columns and  $m$  rows, we can calculate a measure of nestedness among all columns ( $N_{col}$ ) and among all rows ( $N_{row}$ ) by simply averaging all paired values of columns and rows.

Finally, the measure of nestedness for the whole matrix is given by:

$$NODF = \frac{\sum N_{paired}}{\left(\frac{n(n-1)}{2}\right) + \left(\frac{m(m-1)}{2}\right)}$$

In the context of studies of individual specialization, one form of diet variation arises when individuals differ in their niche breadth, such that some individuals diet is a subset of other individuals' diet. This is revealed by a nestedness metric, which may be large (indicating nesting) or small (indicating clustering).

## Value

The result is a list of class 'RInSp' composed of:

NODF	Value of the index of nestedness.
Nrows	Value of the index of nestedness for rows.
Ncols	Value of the index of nestedness for columns.
R	Binary matrix with individuals as rows and resources as columns. This matrix can be imported into the software PAJEK ( <a href="http://vlado.fmf.uni-lj.si/pub/networks/pajek/">http://vlado.fmf.uni-lj.si/pub/networks/pajek/</a> ) to draw a binary bipartite network of diet connectance between individuals (one set of nodes) and resources (a second set of nodes).
NpR	The degree of nestedness observed for rows.
NpC	The degree of nestedness observed for columns.

## Author(s)

Dr. Nicola ZACCARELLI

## References

Almeida-Neto M., Guimaraes P., Guimaraes P.R., Loyola R.D., and Ulrich, W. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**: 1227-1239.

## See Also

For further indices see the R package “vegan”. Please consider that the NODF implementation lacks the ordering and weighting options of *nestednodf* available in “vegan”.

## Examples

```
# NODF example with stickleback data from Bolnick and Paull 2009
data(Stickleback)
# Select a single spatial sampling site (site D)
SiteD = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13),
subset.rows = c("Site", "D"))
Nesting = NODF(SiteD)
rm(list=ls(all=TRUE))
```

---

Null.Hp.RInSp

---

*Generate a data set based on a specific null hypothesis*


---

## Description

The procedure let you build an integer/count dataset of class RInSp:

- by fixing the number of items/resources per individual and the population diet;
- by providing a specific dataset of class RInSp and an option of “sum” or “average” for the population diet.

It builds a table with a specified number of individuals and for each, a resources vector is created using a multinomial distribution with probability equal to the specified population diet and number of items.

As the procedure checks for zero sum columns/rows, it can be time consuming.

## Usage

```
Null.Hp.RInSp(dataset, prop = "sum")
```

## Arguments

- |         |   |
|---------|---|
| dataset | Object of class RInSp with data of type “integer” or a numeric integer vector specifying the number of items for each individual in the simulated data set.   |
| prop    | Option specifying the type of calculation for population diet. Two types are valid when “dataset” is an object of class RInSp: “sum” or “average”. For details see <a href="#">pop.diet</a> . Or a numeric vector when the counts of items for individuals is provided. |

## Details

The procedure can take a long time as it assure that the final data set has no columns/rows sum equal to zero.

It can be used under two scenarios:

- when a specific data set of integer/count data is of interest, it can be used as input to generate a null data set with the same number of individuals/resources and total number of diet items, but the probability of assignment can be derived by the population diet (providing the “sum” or “average”) or a diet specified by the user;

- to generate a null data set by specifying the total number of items for each individual and the probability of the resource type in the diet using a diet provided.

When the user provides the number of items for individuals, the procedure checks for negative values and it converts the numbers into integers.

When the diet is specified by the user, the procedure will convert all values into positive numbers and recalculate the proportions of each entry, so to ensure to have a vector of probabilities (i.e., sum equal to one, values between zero and one).

A progress bar shows the progress in generating random samples till a maximum of 400 new random data sets.

The procedure can easily create a big three dimensional array. This object can create problems as it can consume an huge amount of memory. If needed use a package like “ff” and a for loop to create an array able to hold the results.

## Value

An object of class RInSp holding the null model data set.

## Author(s)

Dr. Nicola ZACCARELLI and Dr. I. Bolnick

## See Also

Function [pop.diet](#).

## Examples

```
# Example with stickleback data from Bolnick and Paull 2009
data(Stickleback)
# Null model using a specified data set
SiteA = import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "A"))
Eresult = Eindex(SiteA, index = "saramaki", jackknife = FALSE)
Nulldata = Null.Hp.RInSp(SiteA, prop="average")
Eresult = Eindex(Nulldata, index = "saramaki", jackknife = FALSE)

# Null model with diet and per individual resource
# Generate a number-of-items-per-individual vector for 30 individuals
ind.resources = floor(runif(30, 15, 120))
# Generate a population diet with 6 items
```

```
diet = c(10, 30, 15, 60, 100)
Nulldata2 = Null.Hp.RInSp(ind.resources, prop= diet)
Eresult2 = Eindex(Nulldata2, index = "saramaki", jackknife = FALSE)
rm(list=ls(all=TRUE))
```

---

overlap	<i>Mean pairwise overlap</i>
---------	------------------------------

---

### Description

This procedure calculates the pairwise diet overlap between all individuals in a sample.

### Usage

```
overlap(dataset)
```

### Arguments

dataset            Object of class RInSp with data.

### Details

For a sample size of  $N$  individuals, this will produce an  $N * N$  matrix, each cell  $o_{ik}$  representing the diet overlap between individual  $i$  and individual  $k$ . The diagonal is all ones because an individual has a 100% overlap with itself. Pairwise overlap is calculated by determining the proportional similarity between each individual.

The script reports the average pairwise diet overlap (not including diagonal values), the average pairwise diet dissimilarity (i.e., one minus the average pairwise diet overlap) and the complete diet overlap matrix. The overlap matrix can be useful for testing whether diet (dis)similarity is a function of various other metrics of between-individual difference or similarity (e.g., Bolnick and Paull 2009).

### Value

Return an matrix of three columns:

meanoverlap	The mean pairwise overlap for all pairs of individuals, excluding the diagonal.
meanindividualoverlap	The mean pairwise overlap for individual $i$ and all others, excluding the diagonal.
meandissimilarity	The mean pairwise dissimilarity overlap for all pairs of individuals, excluding the diagonal.
overlapmatrix	The probability associated to the likelihood value.
parameter=0	Internal option for plotting results.

**Author(s)**

Dr. Nicola ZACCARELLI

**References**

Bolnick, D.I., and Paull, J.S., 2009. Morphological and dietary differences between individuals are weakly but positively correlated within a population of threespine Stickleback. *Evolutionary Ecology Research* **11**: 1217-1233.

Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.

**Examples**

```

data(Stickleback)
# Select a single spatial sampling site (site D)
GutContents_SiteD = import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "D"))
similarity = overlap(GutContents_SiteD)

## Not run:
# Generate a plot similar to Bolnick and Paull 2009, testing whether
# morphological differences between individuals are positively related
# to diet differences (or conversely are negatively related to diet similarity)
GutContents = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13))
similarity = overlap(GutContents)
diet.similarity = as.dist(similarity$overlapmatrix)
size.dissimilarity = dist(GutContents$info[,3])
plot(diet.similarity ~ size.dissimilarity, xlab = "Size dissimilarity",
ylab = "Diet similarity", pch = 16, cex = 0.4)
model = lm(diet.similarity ~ size.dissimilarity)
abline(model, lwd = 3, col = "red")

# The ade4 package is needed to complete the analysis
library(ade4)
mantel.rtest(diet.similarity, size.dissimilarity)

## End(Not run)

```

---

pop.diet

*Calculate population diet*


---

**Description**

This function calculates the population diet using two different approaches: the “sum” or the “average” method.

**Usage**

```
pop.diet(dataset, prop = "sum")
```

**Arguments**

dataset	Object of class RInSp with data.
prop	Option specifying the type of calculation. Two types: “sum” or “average”.

**Details**

The “sum” approach is the most straightforward. One sums up all resource counts within a category across all individuals to get the population’s use, then determine the proportion of each resource category in the population’s repertoire. The proportion  $q_j$  of the resource  $j$  in the population’s diet is:

$$q_j = \frac{\sum_i n_{ij}}{\sum_{j_i} n_{ij}}$$

The drawback of this approach is that individuals that eat large numbers of items, or larger total mass of items, will bias the population to look more like them.

The “average” method (average proportion) circumvents this problem by first converting individual diets into proportions  $p_{ik}$ , then averaging these proportions for each resource  $k$ .

Along with the population’s diet the procedure calculates the Levins’ D index (Levins 1968) of diversity as:

$$D = 1 - \frac{1}{\sum q_j^2}$$

**Value**

Returns a list of information:

popdiet	Vector of the specified population diet.
popdtype	The type of population diet.
richness	The richness of resources (for now it has sense only for integer data type).
D	The Levins’ D index for resources.

**Author(s)**

Dr. Nicola ZACCARELLI

**References**

Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.

Levins R. 1968. *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton Univ. Press.

**See Also**

Function [Null.Hp.RInSp](#).

**Examples**

```

data(Stickleback)
# Select a single spatial sampling site (site D)
GutContents_SiteD = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13),
subset.rows = c("Site", "D"))
PopDiet = pop.diet(GutContents_SiteD, prop = "average")
rm(list=ls(all=TRUE))

```

PSicalc

*Calculate the proportional similarity index***Description**

A measure of individual specialization proposed by Bolnick et al. (2003), based on the average pairwise overlap of the niche distribution of individuals and the population.

**Usage**

```
PSicalc(dataset, pop.diet = "sum", exclude = FALSE, replicates=999, precision = 1e-9)
```

**Arguments**

dataset	Object of class RInSp with data.
pop.diet	Option specifying the way to calculate the population diet. Valid values are "sum" or "average".
exclude	A logical value to exclude or not the individual in variance estimates. Defalut is FALSE, i.e. keep all individuals.
replicates	Number of replicates for Monte Carlo resampling.
precision	a filter to remove from the PS matrix all values which are lower than "precision". In this way some interactions are removed due to their small weight in the PS matrix.

**Details**

Both Schoener (1968) and Feinsinger et al. (1981) advocated using Czekanowski's proportional similarity index (PS) for interspecific niche studies. Adapted to individual-level analyses, the diet overlap between an individual and the population is:

$$PS_i = 1 - 0.5 \sum_j |p_{ij} - q_j|$$

where  $p_{ij}$  is the frequency of category  $j$  in the individual  $i$ 's diet, and  $q_j$  is the frequency of category  $j$  in the population as a whole. The prevalence of individual specialization (IS) in the population is then measured by the average of individuals' PS values:

$$IS = \frac{\sum_i (PS_i)}{N}$$

Monte Carlo resampling for hypothesis testing is implemented for the case when all individuals are kept.

The user has the option of calculating each individual's  $PS_i$  and  $Var(PS_i)$  using  $q_j$  values calculated by either including or excluding individual  $i$ .

### Value

The function returns a list composed by:

PSi	A column vector with the individual proportional similarity index for each individual $i$ .
IS	The value of the prevalence of individual specialization.
PSi.montecarlo	Monte Carlo resampling simulation values for $PS_i$
Var.montecarlo	Monte Carlo resampling simulation values for $PS_i$ 's variance
VarPSi	A column vector of the estimated variance for each individual.
population.diet	A vector describing the population diet. A matrix is returned when "exclude" is TRUE.
IS.pvalue	Probability value for the observed IS against the Monte Carlo resampling simulation. Consider to use more than 1000 simulations.
montecarlo	Monte Carlo resampling simulation values for $PS_i$
num.individuals	The number of individuals of the data set.
parameter	The value points to the column of the "montecarlo" matrix for plotting results.
Precision	The value of the selected threshold for precision.

### Author(s)

Dr. Nicola ZACCARELLI

### References

- Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.
- Feinsinger, P., Spears, E. E., and Poole, R. W. 1981. A simple measure of niche breadth. *Ecology* **62**: 27-32.
- Schoener, T.W. 1968. The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49**: 704-726.

### Examples

```
# PSicalc example using stickleback data from Bolnick and Paull 2009
data(Stickleback)
# Select a single spatial sampling site (site A)
GutContents_SiteA <- import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "A"))
# Warning, the number of replicates is set low to speed up he example!
```



```
# Note, for real analyses we recommend to set replicates => 999
PSi <- PSicalc(GutContents_SiteA, exclude = FALSE, replicates = 99)
sumMC.RInSp(PSi)
rm(list=ls(all=TRUE))
```

---

RInSp2vegan

*Vegan support procedure*

---

### Description

The procedure converts an RInSp class object into a dataframe to be used as input for the “Vegan” package to calculate nestedness indices.

### Usage

```
RInSp2vegan(dataset)
```

### Arguments

dataset            Object of class RInSp with data of type “double”, “integer” or “proportions”.

### Value

The result is a binary dataframe with rows and columns names from the input object.

### Author(s)

Dr. Nicola ZACCARELLI

### See Also

For further indices see the R package “vegan”.

### Examples

```
data(Stickleback)
# Select a single spatial sampling site (site D)
GutContents = import.RInSp(Stickleback, row.names = 1, info.cols = c(2:13))
risul = RInSp2vegan(GutContents)
rm(list=ls(all=TRUE))
```

---

Stickleback

*Example of data from threespine Stickleback gut contents*

---

## Description

Example of a raw data matrix to be used with the `import.RInSp` procedure. Data are from Bolnick and Paull (2009), studying individual specialization in threespine Stickleback (*Gasterosteus aculeatus*) from 5 sites within one lake in British Columbia. Data columns include:

- FishID: a unique identifier for each individual fish;
- Site: A-E, five locations around the lake where fish were collected;
- Mass: fish mass in grams;
- Standard.length, Body.width, gape.width: fish dimensions in centimetres;
- raker.number: gill raker number;
- raker.length: gill raker length (mm);
- Sex: sex;
- MicrogN, DeltaAir, MicrogC, DeltaPDB: stable isotope data for C and N stable isotope ratios.

All other columns are counts of prey categories in Stickleback stomach contents.

## Usage

```
data(Stickleback)
```

## References

Bolnick, D.I., and Paull, J.S. 2009. Morphological and dietary differences between individuals are weakly but positively correlated within a population of threespine Stickleback. *Evolutionary Ecology Research* **11**: 1217-1233.

## Examples

```
data(Stickleback)
summary(Stickleback)
rm(list=ls(all=TRUE))
```

---

`sumMC.RInSp`*Summary function for Monte Carlo resampling results*

---

## Description

The function provides summary statistics on the Monte Carlo resampling simulated index distributions. It plots an histogram with the position of the actual value of the index for the original dataset (blue line) and the 2.5% and 97.5% percentile of the simulated distribution (red dashed lines).

It works for objects of class 'RInSp'.

## Usage

```
sumMC.RInSp(dataset)
```

## Arguments

`dataset`            Dataset of class RInSp.

## Author(s)

Dr. Nicola ZACCARELLI

## See Also

See also [WTcMC](#), [WTdMC](#), or [Emc](#).

## Examples

```
# sumMC.RInSp example using stickleback data from Bolnick and Paull 2009
data(Stickleback)
# Select a single spatial sampling site (site A)
GutContents_SiteA = import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "A"))
EmcRes = Emc(GutContents_SiteA, replicates=499)
sumMC.RInSp(EmcRes)
rm(list=ls(all=TRUE))
```

---

 Trout
 

---



---

*Example of continuous data from fish prey lengths of brown trout*


---

### Description

The example reproduces data published in Kahilainen & Lehtonen (2001) on prey lengths from a sample of 59 individuals of stocked (S) and native (N) brown trout (*Salmo trutta*) in the subarctic Lake Muddusjarvi in northern Finland.

Column names refer to:

- FishID: a unique identifier for each individual fish within the geographic area;
- Type: stocked (S) and native (N) brown trout.

All other columns are lengths (in mm) of fish prey found in stomachs.

### Usage

```
data(Trout)
```

### References

Kahilainen, K. and Lehtonen, H. 2001. Resource use of native and stocked brown trout *Salmo trutta* L., in a subarctic lake. *Fisheries Management and Ecology* **8**: 83-94.

### Examples

```
# Summary of total length of preys in stomach content
# by geographic region
data(Trout)
# Mean prey length by individual
# First change zeros to NA
troutTMP = Trout
troutTMP[ troutTMP == 0 ] = NA
individuals = by(troutTMP[, 3:7], troutTMP[, 2], rowMeans, na.rm=TRUE)
TroutN = subset(Trout, Type == "N", PL1:PL5)
TroutN = TroutN[TroutN > 0]
TroutS = subset(Trout, Type == "S", PL1:PL5)
TroutS = TroutN[TroutS > 0]
boxplot(c(TroutS, TroutN) ~ c(rep("N", length(TroutN)), rep("S", length(TroutS))))
rm(list=ls(all=TRUE))
```

**Description**

The program calculates the Total Niche Width (TNW), and breaks TNW down into its Between Individual Component (BIC) and Within Individual Component (WIC). It reports these three statistics and the proportion WIC/TNW. Niche widths of the population, within individuals, and among individuals are calculated following Roughgarden's formulas (1974).

A Monte Carlo resampling routine is used to calculate a null distribution for each of these statistics and generate a p-value testing the null hypothesis that all individuals sample equally from the population diet distribution.

**Usage**

```
WTcMC(dataset, replicates = 999, weight = "equal", print.ris=TRUE)
```

**Arguments**

dataset	Object of class RInSp with data of type "double".
replicates	Number of replicates for the Monte Carlo resampling.
weight	The weighting scheme for individuals. Default value is "equal". Alternative value is "N_items".
print.ris	Printing results to screen. Default is TRUE.

**Details**

Roughgarden (1972) suggested that the total niche width of a population (TNW) can be broken down into two components: the variation in resource use within individuals (within-individual component, WIC), and the variance between individuals (between-individual component, BIC) so that  $TNW = WIC + BIC$ . Originally developed for within- and between-phenotype (or between-sex) variation, it can be applied at the individual level. Assuming that variation in niche parameters can be expressed along a single continuous dimension  $x$  (e.g., prey size), let  $X$  be a matrix of diet data where each element  $x_{ij}$  is the size (or other measure) of the  $j$ th prey item in individual  $i$ 's diet. Then,

$$TNW = Var(x_{ij})$$

$$WIC = E(Var(x_j|i))$$

$$BIC = Var(E(x_j|i))$$

The relative degree of individual specialization can be measured as the proportion of TNW explained by within-individual variation, WIC/TNW. As this value approaches 1, all individuals utilize

the full range of the population's niche, whereas smaller values indicate decreasing inter-individual overlap and hence higher individual specialization.

Roughgarden's WIC/TNW is limited to continuous diet data. To carry this approach over to discrete data such as the frequency of alternate prey taxa in the diet, Roughgarden (1979:510) proposed a measure that uses the Shannon-Weaver index as a proxy for variance (See WTdMC for details).

In case of a marked presence of individuals with one or few prey items, the relation  $TNW = BIC + WIC$  may not hold. In this case there is the option to weight the number of items per individuals following two alternatives:

- "equal": weighting by the number of items in each individual's diet, so those with more data contribute more to estimating parameters;
- "N\_items": weighting each individual equally regardless of diet items number; requires weighting each diet item by the inverse of the number of items in the individual's diet for calculating TNW.

The Monte Carlo resampling procedure first determines the number of prey used by each individual,  $n_i$ . Subsequently, the procedure randomly reassigns each individual  $n_i$  prey items drawn from the population distribution of items. This is repeated a user-specified number of times. Hence, the resampled population yields a null model corresponding to a population composed of generalists that sample randomly from the population's diet, and have diet sizes equal to those of the observed data set. The original statistics are merged into the Monte Carlo set.

## Value

The result is a list of class "RInSp" composed of:

WonT	The value of WIC/TNW for the dataset "dataset".
p.value	The Monte Carlo resampling probability value for the calculated WIC/TNW ratio.
montecarlo	is a matrix with the replicate number 'REP', the value of 'WIC', the value of 'BIC', the value of 'TNW' and the derived value for 'WonT'. The first line holds the values for the provided dataset, and the following lines the results of the Monte Carlo resampling.
weight	The weighting scheme for individuals used.
parameter	An index used in <code>sumMC.RInSp</code> to point to the column of the matrix 'montecarlo' for plotting results.

## Author(s)

Dr. Nicola ZACCARELLI. Please note that the C code is partially based on the C++ source of IndSpec from Bolnick et al. (2002). For the "weight" option Travis Ingram is acknowledged (Ingram et al. 2011).

The original source code of IndSpec1.0 is available from the "Ecological Archives" of the Ecological Society of America (<http://esapubs.org/archive/>): identifier E083-056-S1; <http://esapubs.org/archive/eco1/E083/056/>.

## References

Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.

Ingram, T., Stutz, W.E., Bolnick, D.I. 2011. Does intraspecific size variation in a predator affect its diet diversity and top-down control of prey? *PLoS ONE* **6**: e20782.

Roughgarden, J. 1974. Niche width: biogeographic patterns among *Anolis lizard* populations. *American Naturalist* **108**: 429-411.

### See Also

Function [sumMC.RInSp](#).

### Examples

```
# Example of analysis with the trout data from Finland
data(Trout)
TroutRIS= import.RInSp(Trout, col.header=TRUE, row.names=1, info.cols=2, data.type="double")
# Warning, the number of replicates is set low to speed up the example!
# Note, for real analyses we recommend to set replicates => 999
TroutRIS_MC = WTdMC(TroutRIS, replicates = 99)
sumMC.RInSp(TroutRIS_MC)
rm(list=ls(all=TRUE))
```

---

WTdMC

*Monte Carlo resampling of WIC/TNW for the discrete case*

---

### Description

The program calculates the Total Niche Width (TNW), and breaks TNW down into its Between Individual Component (BIC) and Within Individual Component (WIC). It reports these three statistics and the proportion WIC/TNW. Niche widths of the population, within individuals, and among individuals are calculated using the Shannon-Weaver diversity index on count data, following Roughgarden's formulas (1974) and Bolnick et al. (2002). A Monte Carlo resampling routine is used to calculate a null distribution for each of these statistics and generate a p-value testing the null hypothesis that all individuals sample equally from the population diet distribution. Note that the bootstrapping procedure is only biologically meaningful when applied to integer data representing counts of individual prey within individual predators' diets, and should not be applied to other forms of data such as prey mass or proportion of total volume.

### Usage

```
WTdMC(dataset, pop.diet = "sum", replicates=999, print.ris=TRUE)
```

### Arguments

dataset	Object of class RInSp with data of type "integer".
pop.diet	Option specifying the way to calculate the population diet. Valid values are "sum" or "average". See <a href="#">pop.diet</a> for details.
replicates	Number of replicates for the Monte Carlo resampling.
print.ris	Printing results to screen. Default is TRUE.

## Details

The Roughgarden's measures are the most widely reported indices of diet variation, due to historical precedent. They are closely analogous to an Analysis of Variance, in that one partitions the total variation in resource use (TNW) into within- and between-individual components (WIC and BIC). An advantage of this approach is that one can simultaneously evaluate the degree of diet variation (WIC/TNW), and shifts in individual and population niche widths. A drawback is that, for categorical diet data, the Shannon-Weaver measure of diversity can lead to biased estimates of diet variation (Bolnick et al. 2002). Despite its historical precedent, users are advised to avoid using WIC/TNW. This bias is not a concern for WIC/TNW measured on quantitative traits (prey size) rather than prey categories, but this is rarely done.

Roughgarden (1979, pg 510) proposed a measure to describe the niche width based on Shannon-Weaver information theory. The following formulae are equivalent to Roughgarden's formulation, although we use different notation. Let  $N$  be a matrix of diet data, where elements  $n_{ij}$  represent the number (or mass) of diet items in individual  $i$ 's diet that fall in category  $j$ . This raw data matrix is then transformed into a proportion matrix  $P$ , with elements  $p_{ij}$  describing the proportion of the  $j$ th resource category in individual (or phenotype)  $i$ 's diet.

Using these proportions or the raw numerical counts one then calculates population diet proportions ( $q_k$ , see [pop.diet](#) for details). These data can then be used to calculate WIC, BIC, and TNW (Roughgarden 1979; Bolnick et al. 2002):

$$WIC = \sum_i p_i \left( - \sum_k p_{ik} \ln(p_{ik}) \right)$$

$$BIC = \sum_k p_i \ln(p_i) - \sum_k q_k \left( - \sum_i t_{ik} \ln(t_{ik}) \right)$$

$$TNW = - \sum_k q_k \ln(q_k)$$

where:  $p_i$  is the proportion of all resources used by individual  $i$  ( $\frac{\sum_k n_{ik}}{\sum_i \sum_k n_{ik}}$ );  $q_k$  is the proportion of the  $k$ th resource category in the population's niche, and  $t_{ik}$  is the proportion of the population's total use of resource  $k$  that was used by individual  $i$  (i.e.,  $\frac{n_{ik}}{\sum_i n_{ik}}$ ).

It is important to remember that  $TNW = BIC + WIC$ , and one can easily calculate the proportion WIC, which we refer to as WIC/TNW.

Be aware that monophagous individuals (100% of diet from a single resource) have excessively large leverage drawing WIC towards zero, possibly leading to downward-biased WIC/TNW exaggerating individual specialization. The procedure prints a warning message with the name of monophagous individuals in the dataset.

The Monte Carlo resampling procedure first determines the number of prey used by each individual,  $n_i$ . It then determines the frequencies with which the population as a whole used the various resources (see [pop.diet](#) for details). Subsequently, the procedure randomly reassigns each individual  $n_i$  prey drawn from the population frequency distribution, and then recalculates a null measure of diet variation. This is repeated a user-specified number of times. Hence, the resampled population



yields a null model corresponding to a population composed of generalists that sample randomly from the population's diet, and have diet sizes equal to those of the observed data set. The original statistics are merged into the Monte Carlo set.

### Value

The result is a list of class 'RInSp' composed of:

WonT	The value of WIC/TNW for the dataset "dataset".
Zeros	A character vector with the names of individuals with Shannon-Weaver scores equal to zero. Or zero otherwise.
p.value	The Monte Carlo resampling probability value for the calculated WIC/TNW ratio;
montecarlo	A matrix with the replicate number 'Zeroes', the value of 'WIC', the value of 'BIC', the value of 'TNW' and the derived value for 'WonT'. The first line holds the values for the provided dataset, and the following lines the results of the Monte Carlo resampling.
parameter	An index used in <code>sumMC.RInSp</code> to point to the column of the matrix 'montecarlo' for plotting results.

### Author(s)

Dr. Nicola ZACCARELLI. Please note that the C code is partially based on the C++ source of IndSpec from Bolnick et al. (2002).

The original source code of IndSpec1.0.exe is available from the "Ecological Archives" of the Ecological Society of America (<http://esapubs.org/archive/>): identifier E083-056-S1; <http://esapubs.org/archive/eco1/E083/056/>.

### References

Bolnick, D.I., L.H. Yang, J.A. Fordyce, J.M. Davis, and Svanback, R. 2002. Measuring individual-level resource specialization. *Ecology* **83**: 2936-2941.

Roughgarden, J. 1979. Theory of population genetics and evolutionary ecology: an introduction. Macmillan, New York, NY, USA.

### See Also

For the continuous case see [WTcMC](#) and for a summary statistics the function `sumMC.RInSp`.

### Examples

```
data(Stickleback)
# Select a single spatial sampling site (site A)
GutContents_SiteA = import.RInSp(Stickleback, row.names = 1,
info.cols = c(2:13), subset.rows = c("Site", "A"))
# Warning, the number of replicates is set low to speed up the example!
# Note, for real analyses we recommend to set replicates => 999
WT = WTdMC(GutContents_SiteA, replicates = 99)
sumMC.RInSp(WT)
rm(list=ls(all=TRUE))
```

# Index

## \*Topic **dataset**

Stickleback, [26](#)  
Trout, [28](#)

## \*Topic **function**

import.RInSp, [11](#)  
Likelihood and Wi, [13](#)  
overlap, [20](#)  
pop.diet, [21](#)  
PSicalc, [23](#)  
sumMC.RInSp, [27](#)

## \*Topic **models**

Eindex, [4](#)  
Emc, [7](#)  
Hier2L, [10](#)  
MCp.RInSp, [15](#)  
NODF, [16](#)  
Null.Hp.RInSp, [18](#)  
RInSp2vegan, [25](#)  
WTcMC, [29](#)  
WTdMC, [31](#)

## \*Topic **univar**

Eindex, [4](#)  
Emc, [7](#)  
Hier2L, [10](#)  
MCp.RInSp, [15](#)  
NODF, [16](#)  
Null.Hp.RInSp, [18](#)  
RInSp2vegan, [25](#)  
WTcMC, [29](#)  
WTdMC, [31](#)

## \*Topic **utilities**

import.RInSp, [11](#)  
Likelihood and Wi, [13](#)  
overlap, [20](#)  
pop.diet, [21](#)  
PSicalc, [23](#)  
sumMC.RInSp, [27](#)

Eindex, [2](#), [3](#), [4](#), [7](#), [9](#)

Emc, [2](#), [3](#), [7](#), [7](#), [15](#), [16](#), [27](#)

Hier2L, [10](#)

import.RInSp, [7](#), [11](#), [26](#)

like.Wi, [2](#)

like.Wi (Likelihood and Wi), [13](#)

Likelihood and Wi, [13](#)

MCp.RInSp, [15](#)

NODF, [16](#)

Null.Hp.RInSp, [9](#), [18](#), [22](#)

overlap, [20](#)

pop.diet, [7](#), [15](#), [16](#), [18](#), [19](#), [21](#), [31](#), [32](#)

PSicalc, [3](#), [15](#), [16](#), [23](#)

RInSp (RInSp-package), [2](#)

RInSp-package, [2](#)

RInSp2vegan, [25](#)

Stickleback, [26](#)

sumMC.RInSp, [9](#), [27](#), [30](#), [31](#), [33](#)

Trout, [28](#)

WTcMC, [10](#), [11](#), [15](#), [16](#), [27](#), [29](#), [33](#)

WTdMC, [15](#), [16](#), [27](#), [30](#), [31](#)