

Package ‘samplesim’

May 12, 2017

Type Package

Title Sample size effects in stable isotope mixing solutions

Version 1.0

Date 2014-09-27

Author Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

Maintainer Nicolas Casajus <nicolas.casajus@gmail.com>

Description This package allows the investigation of the effect of sample size on estimates and precision of stable isotope mixing solutions calculated with the package `siar` (written by Andrew Parnell et al.). Sample sizes are modified assuming a normal distribution with a user defined mean and standard deviation. Samples of different sizes are created from this distribution, and mixing proportions are estimated for several replicates of each sample size using the function `siarmcmcdirchletv4` (package `siar`) and default values for the MCMC. Authors thank Rolf Ims for his contribution.

License GPL-2

Depends `siar`, `plyr`, `ggplot2`

Suggests `knitr`, `rmarkdown`

VignetteBuilder `knitr`

NeedsCompilation no

RemoteType `github`

RemoteHost <https://api.github.com>

RemoteRepo `samplesim`

RemoteUsername `ahasverus`

RemoteRef `master`

RemoteSha `459cc8e1e50eee4aa0963948fcd289652bbebbac`

GithubRepo `samplesim`

GithubUsername `ahasverus`

GithubRef `master`

GithubSHA1 `459cc8e1e50eee4aa0963948fcd289652bbebbac`

R topics documented:

samplesim-package	2
credint	3
credintt	3
dataConsumer	4
dataPrey	4
format_sources	5
get_output	6
meansd.nd	8
plot_samplesim	8
rawPrey	9
samplesim	10

Index	13
-------	----

samplesim-package	<i>Sample size effects in stable isotope mixing solutions</i>
-------------------	---

Description

This package allows the investigation of the effect of sample size on estimates and precision of stable isotope mixing solutions calculated with the package [siar](#) (written by Andrew Parnell et al.). Sample sizes are modified assuming a normal distribution with a user defined mean and standard deviation. Samples of different sizes are created from this distribution, and mixing proportions are estimated for several replicates of each sample size using the function `siarmcmcdirchletv4` (package `siar`) and default values for the MCMC.

Details

Package: samplesim
Type: Package
Version: 1.0
Date: 2017-05-11
License: GPL (>= 2)

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich
Maintainer: Nicolas Casajus <nicolas.casajus@gmail.com>

References

Lecomte N., Ehrich D., Casajus N., Berteaux D., Giroux M.-A., Yoccoz N.G. How many is enough? An R package for evaluating the effect of sample size on estimates and precision of stable isotope mixing solutions. Submitted in *Methods in Ecology and Evolution*.

See Also[siar](#)

`credint`*samplesim internal function*

Description

`samplesim` internal function.

Details

This function is a low level function. Not for a direct use.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

See Also[samplesim](#)

`credintt`*samplesim internal function*

Description

`samplesim` internal function.

Details

This function is a low level function. Not for a direct use.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

See Also[samplesim](#)

dataConsumer	<i>Consumer plasma data</i>
--------------	-----------------------------

Description

A 2 column and 60 row data frame which contains the isotopic plasma values of 60 consumers over 2 isotopes. Data come from Killengreen et al. (2011).

Usage

```
data(dataConsumer)
```

Format

A data frame with 60 observations on the following 2 variables:

d13C Plasma values of the d13C isotope

d15N Plasma values of the d15N isotope

Details

This dataset is designed for a direct use by the function [samplesim](#).

References

Killengreen S., Lecomte N., Ehrich D., Schott T., Yoccoz N.G. and Ims R.A. (2011) The importance of marine vs. human-induced subsidies in the maintenance of an expanding mesocarnivore in the arctic tundra. *Journal of Animal Ecology*, 80, 1049-1060.

Examples

```
data(dataConsumer)
head(dataConsumer)
```

dataPrey	<i>Preys plasma data</i>
----------	--------------------------

Description

A 5 column and 6 row data frame which contains the isotopic plasma values of 6 preys (sources) over mean and standard deviation values of 2 isotopes. Data come from Killengreen et al. (2011).

Usage

```
data(dataPrey)
```

Format

A data frame with 6 observations on the following 5 variables:

Sources Name of sources
 Meand13C Mean values of the d13C isotope
 SDd13C Standard deviations of the d13C isotope
 Meand15N Mean values of the d15N isotope
 SDd15N Standard deviations of the d15N isotope

Details

This dataset is designed for a direct use by the function `samplesim`.

References

Killengreen S., Lecomte N., Ehrich D., Schott T., Yoccoz N.G. and Ims R.A. (2011) The importance of marine vs. human-induced subsidies in the maintenance of an expanding mesocarnivore in the arctic tundra. *Journal of Animal Ecology*, 80, 1049-1060.

Examples

```
data(dataPrey)
print(dataPrey)
```

format_sources

Format source data for stable isotope analysis

Description

This function computes mean and standard deviation of isotope concentrations from sources raw measurements.

Usage

```
format_sources(data, labels)
```

Arguments

data	A data frame containing the isotope values (column) for each individual prey (row). The first column must contain the source name. See <code>rawPrey</code> for an example.
labels	A list containing the different source names where the name of each element of the list has to be specified. Each element of the list can contain more than one source name in case of grouping. See the example below.

Details

This function works with two or more than two isotopes.

Value

A data frame with each source as a separate row and (2 x number of isotopes) columns. An additional column (first) contains the source names. For each isotope mean and standard deviation are returned.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

See Also

[rawPrey](#), [samplesim](#)

Examples

```
# Import example dataset
data(rawPrey)

# Group similar preys
src <- list(
  Marine    = c("Ristri", "Urilom"),
  Reindeer  = c("Rantar"),
  Voles     = c("Micoec"),
  Lemming   = c("Lemlem"),
  Ptarmigan = c("Laglag", "Lagmut"),
  Birds     = c("Melnig"))

# Format data
(meansdiso <- format_sources(rawPrey, labels = src))
```

get_output

Import samplesim simulation results

Description

This function imports samplesim simulation results. More specifically it loads the median and width of credible intervals.

Usage

```
get_output(name = "simulation_1", change = FALSE, reference = NULL)
```

Arguments

name	Name of the simulation to plot results.
change	If TRUE, values are expressed as a percentage of change.
reference	The reference state to compute percentage of change. Default is the minimum value of sample size.

Value

A data frame with five columns:

replicate	The number of replicate.
source	The source name.
size	The sample size.
value	The value of medians/widths of credible intervals.
type	The label of value (i.e. medians or widths)

If 'change' is TRUE, then the column replicate is omitted and results are aggregated over replicates.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

References

Lecomte N., Ehrich D., Casajus N., Berteaux D., Giroux M.-A., Yoccoz N.G. How many is enough? An R package for evaluating the effect of sample size on estimates and precision of stable isotope mixing solutions. Submitted in Methods in Ecology and Evolution.

See Also

[samplesim](#), [plot_samplesim](#)

Examples

```
# Load datasets
data(dataPrey)
data(dataConsumer)

# Effect of sample size by modifying source 4
samplesim(
  dataConsumer,
  dataPrey,
  type = "one source",
  nsamples = c(5, 10, 25),
  nrep = 10,
  modwhich = 4,
  name = "sim_3")

# Import results
res <- get_output(name = "sim_3")

# Print results
head(res, 10)

# Import results (percentage of change)
res <- get_output(name = "sim_3", change = TRUE, reference = 5)

# Print results
head(res, 10)
```

meansd.nd	<i>samplesim internal function</i>
-----------	------------------------------------

Description

samplesim internal function.

Details

This function is a low level function. Not for a direct use.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

See Also

[samplesim](#)

plot_samplesim	<i>Plot samplesim simulation results</i>
----------------	--

Description

This plot function is a graphical representation of the effects of sample size on estimates and precision of stable isotope mixing solutions. Two plots are currently available: a plot for the width of the credible interval and one for the median of the posterior distribution, both displayed for each sample size and each source.

Usage

```
plot_samplesim(name = "simulation_1", change = FALSE, reference = NULL)
```

Arguments

name	Name of the simulation to plot results.
change	If TRUE, values are expressed as a percentage of change.
reference	The reference state to compute percentage of change.

Details

This plot function automatically loads results data stored by the [samplesim](#) function from the simulation name. It produces two plots. The first plot presents the width of the credible interval displayed for each sample size and each source. The second represents the median of the posterior distribution for each sample size and each source.

Author(s)

Nicolas Casajus, Nicolas Lecomte, Dorothee Ehrich

See Also

[samplesim](#), [get_output](#)

Examples

```
# Load datasets
data(dataPrey)
data(dataConsumer)

# Effect of sample size by modifying source 4
samplesim(
  dataConsumer,
  dataPrey,
  type = "one source",
  nsamples = c(5, 10, 25),
  nrep = 10,
  modwhich = 4,
  name = "sim_2")

# Visualize results
plot_samplesim(name = "sim_2")

# Visualize results (percentage of change)
plot_samplesim(name = "sim_2", change = TRUE, reference = 5)
```

rawPrey

Raw prey plasma data

Description

A 3 column and 97 row data frame which contains the isotopic plasma values of 97 prey (sources) over 2 isotopes. Data come from Killengreen et al. (2011).

Usage

```
data(rawPrey)
```

Format

A data frame with 97 observations on the following 3 variables:

Species Species code

d13C Plasma values of the d13C isotope

d15N Plasma values of the d15N isotope

Details

This dataset is not designed for a direct use by the function [samplesim](#). The function [format_sources](#) converts these raw data in the samplesim standard (see [dataPrey](#) and the example below).

References

Killengreen S., Lecomte N., Ehrich D., Schott T., Yoccoz N.G. and Ims R.A. (2011) The importance of marine vs. human-induced subsidies in the maintenance of an expanding mesocarnivore in the arctic tundra. *Journal of Animal Ecology*, 80, 1049-1060.

Examples

```
# Import example dataset
data(rawPrey)

# Group similar preys
src <- list(
  Marine    = c("Ristri", "Urilm"),
  Reindeer  = c("Rantar"),
  Voles     = c("Micoec"),
  Lemming   = c("Lemlem"),
  Ptarmigan = c("Laglag", "Lagmut"),
  Birds     = c("Melnig"))

# Format data
(meansdiso <- format_sources(rawPrey, labels = src))
```

samplesim

Effect of sample size in stable isotope mixing models

Description

This function allows investigating the effect of sample size on estimates and precision of stable isotope mixing solutions. User can modify the sample size for one chosen source, for all sources or for the consumer. See details section for further informations.

Usage

```
samplesim(target, sources, type = NULL, nsamples = NULL, modwhich = NULL,
correct = NULL, nrep = 100, interval = 90, name = NULL)
```

Arguments

target	A data frame with consumer isotope values. See example below or dataConsumer for the appropriate format.
sources	A data frame with mean and standard deviation isotope values of sources. See example below or dataPrey for the appropriate format.
type	Character indicating the type of analysis to run. Must be one of 'one source', 'all sources' or 'consumer'.
nsamples	A vector with the sample sizes to simulate.
modwhich	An integer indicating which source has to be modified. This argument has to be specified when type is 'one source'. Otherwise it will be ignored.
correct	Optional. If specified, a data frame with discrimination values. See siar for further details.

nrep	An integer specifying the number of replicates for each sample sizes. Default is 100.
interval	An integer indicating the width of credible interval to use for precision estimation. Default is 90.
name	A character string giving the name of the simulation. If NULL the simulation will be named by the time of the simulation. This name will serve to create a directory in which results will be stored.

Details

This function assesses the sensitivity of isotopes mixing models to variation in numbers of samples from source tissues. This tool can be used prior to full-blown studies in a similar manner than power analyses. It used the function [siarmcmkdirichletv4](#) developed by Andrew Parnell and available in the package [siar](#). User can choose to sample one particular source (argument type sets to 'one source') or all the sources in the same type (argument type sets to 'all sources'). User can also choose to modify consumer data (argument type sets to 'consumer'). Sample sizes are modified assuming a normal distribution with a user defined mean and standard deviation. Samples of different sizes are created from this distribution, and mixing proportions are estimated for several replicates of each sample size with the function [siarmcmkdirichletv4](#).

Value

This function does not return any object in the R console. Results are stored in a directory (argument 'name') and contain four R objects.

intervals	A four dimensions array with the upper and lower bounds of the credible interval for each sample size, replicate and source. First dimension represents lower and upper bounds; second dimension corresponds to the number of sources; third dimension is the number of replicates; and fourth dimension is the number of sample size.
widths	A three dimensions array with the width (precision) of credible intervals for each source, each replicate and each sample size. First dimension corresponds to the number of replicates; second dimension is the number of sources; and third dimension represents the number of sample size.
medians	A three dimensions array with the median (estimate) of credible intervals for each source, each replicate and each sample size. Dimensions are the same as for widths object.
datasets	A four dimensions array with all resampled datasets.

A logfile is also written and contains all parameters of the simulation.

Author(s)

Nicolas Casajush, Nicolas Lecomte, Dorothee Ehrlich

References

Lecomte N., Ehrlich D., Casajush N., Berteaux D., Giroux M.-A., Yoccoz N.G. How many is enough? An R package for evaluating the effect of sample size on estimates and precision of stable isotope mixing solutions. Submitted in *Methods in Ecology and Evolution*.

See Also

[format_sources](#), [plot_samplesim](#), [siarmcmkdirichletv4](#)

Examples

```
# Load datasets
data(dataPrey)
data(dataConsumer)

# Effect of sample size by modifying source 4
samplesim(
  dataConsumer,
  dataPrey,
  type = "one source",
  nsamples = c(5, 10, 25),
  nrep = 10,
  modwhich = 4,
  name = "sim_1")
```

Index

credint, [3](#)
credintt, [3](#)

dataConsumer, [4](#), [10](#)
dataPrey, [4](#), [9](#), [10](#)

format_sources, [5](#), [9](#), [11](#)

get_output, [6](#), [9](#)

meansd.nd, [8](#)

plot_samplesim, [7](#), [8](#), [11](#)

rawPrey, [5](#), [6](#), [9](#)

samplesim, [3–9](#), [10](#)
samplesim-package, [2](#)
siar, [2](#), [3](#), [10](#), [11](#)
siarmcmcdirichletv4, [11](#)