

Package ‘SAMPLE’

September 15, 2026

Type Package

Title Estimate Sampling Effort for Species Occurrence Rates

Version 1.0.1

Maintainer Henrique Bravo <henrique.bravo@rug.nl>

Description Estimates the sampling effort needed to obtain stable species occurrence and prevalence rates from presence-absence data. The method repeatedly subsamples observations, evaluates changes in estimated rates, and reports the minimum stable sample size. For methodological details, see Bravo et al. (2025) <doi:10.1002/ece3.70998>.

URL <https://github.com/yacinebenchehida/SAMPLE>

BugReports <https://github.com/yacinebenchehida/SAMPLE/issues>

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Imports ggplot2 (>= 3.4.2), Rmisc (>= 1.5.1), RColorBrewer (>= 1.1-3), dplyr (>= 1.1.4), magrittr (>= 2.0.3)

Suggests knitr, rmarkdown, testthat (>= 3.0.0), tibble (>= 3.2.1)

NeedsCompilation no

Author Henrique Bravo [aut, cre] (ORCID:
<<https://orcid.org/0000-0003-3086-737X>>),
Yacine Ben Chehida [aut] (ORCID:
<<https://orcid.org/0000-0001-7269-9082>>),
Sancia E. T. van der Meij [aut] (ORCID:
<<https://orcid.org/0000-0002-3759-8945>>)

Repository CRAN

Date/Publication 2026-09-15 11:30:08 UTC

Contents

coral_symbionts	2
plotstab	3
RunPerm	3
SAMPLE	4
stability	5
Index	7

coral_symbionts	<i>Coral Symbionts Data</i>
-----------------	-----------------------------

Description

Example of SAMPLE input. This dataset contains information about coral species and their associated symbionts.

Usage

```
data(coral_symbionts)
```

Format

The first column contains the names of host species, with each row representing a separate sample. Each sample of a single species should have the same name. The remaining columns represent different parasites or symbionts, showing the count of parasites or symbionts present in each individual listed in the first column. If no data is available for an individual, it is denoted as 'NA'.

Source

Bravo, Ben Chehida and van der Meij (2025), supplementary data, [doi:10.1002/ece3.70998](https://doi.org/10.1002/ece3.70998).

References

Bravo, Ben Chehida and van der Meij (2025), [doi:10.1002/ece3.70998](https://doi.org/10.1002/ece3.70998).

Examples

```
# Load the dataset
data(coral_symbionts)
head(coral_symbionts)
```

plotstab

*Plot***Description**

This function generates the prevalence rate plots for all the variables in question (e.g. species, habitats).

Usage

```
plotstab(data, info, outputName = "Results", outputDir = tempdir())
```

Arguments

data	A dataframe generated by the function RunPerm() (a dataframe object).
info	A dataframe generated by the function stability() (a dataframe object).
outputName	Prefix used for the output (a character; default output_N="Results").
outputDir	Output directory (a character). If omitted, files are generated in the R session's temporary directory.

Value

A text file (.txt) with the output values of the analysis, and a PDF file with the generated plots from the analysis.

Examples

```
data("coral_symbionts")
set.seed(812)
perm = RunPerm(input = coral_symbionts, replicates = 50)
stable = stability(data = perm, stability_thresh = 5, success_points = 5, diff = 2)
plotstab(data = perm, info = stable, outputName = "Stability_example")
unlink(file.path(tempdir(), "Stability_example.pdf"))
unlink(file.path(tempdir(), "Stability_example.txt"))
```

RunPerm

*Run permutations***Description**

This function defines the number of permutations (default = 50) and subsequently runs the permutations on an input dataframe. A minimum number of 10 permutations is recommended in order to avoid having very large or very small stability values due to the stochastic nature of the process. On a large dataset, increasing the number of permutations can considerably slow down the analysis.

Usage

```
RunPerm(input, replicates = 50)
```

Arguments

input	Input dataframe (a dataframe object).
replicates	Number of permutation replicates to perform (an integer; default replicates=50).

Value

A dataframe with the output values of the permutation analysis.

Examples

```
data("coral_symbionts")
perm <- RunPerm(input = coral_symbionts, replicates = 50)
perm
```

SAMPLE

Executing this function runs all the steps of the SAMPLE pipeline.

Description

Only the input dataframe is required to run SAMPLE, all the other arguments have default values that can be adjusted (but it is not necessary to do so). To obtain the same output as in the example file, remember to `set.seed(812)` after importing the dataset. Alternatively, SAMPLE can be run using separate functions. To do so, please refer to `RunPerm()` for the first step of this process.

Usage

```
SAMPLE(input, output_N = file.path(tempdir(), "Results"), replicates = 50,
       stability_thresh = 2, success_points = 10, diff = 1)
```

Arguments

input	Input dataframe (a dataframe object).
output_N	Prefix, optionally including a path, used for the output (a character; default is "Results" in the R session's temporary directory). A prefix without a path also writes to the temporary directory.
replicates	Number of permutation replicates to perform (an integer; default replicates=50).
success_points	Number of successive (mean) prevalence rates that are below a threshold (see parameter <code>stability_thresh</code>) used to define stability (an integer; default <code>success_points</code> = 10).
stability_thresh	Threshold used to define stability (an integer). This value will be divided by the square root of the number of replicates (a float; default <code>stability_thresh</code> = 2.0).
diff	Difference between absolute minimum and maximum values among the all the means used to set the stability threshold (a float; default <code>diff</code> = 1.0).

Value

A text file (.txt) with the output values of the analysis, and a PDF file with the generated plots from the analysis.

Examples

```
data("coral_symbionts")
set.seed(812)
output_prefix <- file.path(tempdir(), "Example")
SAMPLE(input = coral_symbionts,
        output_N = output_prefix,
        replicates = 50,
        stability_thresh = 2,
        success_points = 10,
        diff = 1)
unlink(paste0(output_prefix, ".pdf"))
unlink(paste0(output_prefix, ".txt"))
```

stability

Assess the stability

Description

This function determines the stability values for all of the input variables in the dataset (e.g. species, habitats). Stability here is defined as the minimum occurrence value at which the variable in question stayed stable. See the arguments below to understand how they can affect the stability values.

Usage

```
stability(data, stability_thresh = 2, success_points = 50, diff = 1)
```

Arguments

data	A dataframe generated by the function RunPerm().
stability_thresh	Threshold used to define stability (an integer). This value will be divided by the square root of the number of replicates (a float; default stability_thresh = 2.0).
success_points	Number of successive (mean) prevalence rates that are below a threshold (see argument stability_thresh) used to define stability (an integer; default success_points = 50).
diff	Difference between the absolute minimum and maximum values among the all means used to set the stability threshold (a float; default diff = 1.0).

Value

A dataframe with the output values of the stability assessment analysis.

Examples

```
data("coral_symbionts")
set.seed(812)
perm <- RunPerm(input = coral_symbionts, replicates = 50)
stable <- stability(data = perm, stability_thresh = 5 , success_points = 5, diff = 2 )
stable
```

Index

coral_symbionts, [2](#)

plotstab, [3](#)

RunPerm, [3](#)

SAMPLE, [4](#)

stability, [5](#)