Package 'mbRes'

December 18, 2022

Type Package

Title Exploration of Multiple Biomarker Responses using Effect Size

Version 0.1.6

Description Summarize multiple biomarker responses of aquatic organisms to contaminants using Cliff's delta, as described in Pham & Sokolova (2022) <doi:10.1002/ieam.4676>.

Depends R (>= 4.2.0)

Imports stats, ggplot2 (>= 3.4.0), cowplot (>= 1.1.1), magrittr (>= 2.0.3), tibble (>= 3.1.8), dplyr (>= 1.0.10), forcats (>= 0.5.2), tidyr (>= 1.2.1), purrr (>= 0.3.5), data.table (>= 1.14.6), scales (>= 1.2.1)

Suggests RProbSup (>= 3.0)

BugReports https://github.com/phamdn/mbRes/issues

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Author Duy Nghia Pham [aut, cre] (<https://orcid.org/0000-0003-1349-1710>), Inna M. Sokolova [ths] (<https://orcid.org/0000-0002-2068-4302>)

Maintainer Duy Nghia Pham <nghiapham@yandex.com>

Repository CRAN

Date/Publication 2022-12-18 11:10:02 UTC

R topics documented:

`mbRes-package` *mbRes: Exploration of Multiple Biomarker Responses using Effect Size*

Description

Summarize multiple biomarker responses of aquatic organisms to contaminants using Cliff's delta, as described in Pham & Sokolova (2022) doi:10.1002/ieam.4676.

Guidelines

`mbr` and `visual` are the main functions to compute and visualize Cliff's delta and S-value which are results of `cliff` and `resampling`. `setpop`, `simul`, and `plotsam` simulate and visualize a hypothetical dataset. `compare` compares the results of Cliff's delta and two other integrated indices published earlier (i.e., RSI and IBR, see `blaise2002` and `beliaeff2002`). The others (`ggheat` and `ggdot`) are helper functions and are not meant to be called directly by users.

Updates

`mbr.cliff` and `mbr.glass` simply compute and visualize Cliff's delta and Glass's delta.

Copyright

Author(s)

Duy Nghia Pham & Inna M. Sokolova

`beliaeff2002` *Compute Integrated Biomarker Index*

Description

`beliaeff2002` calculates IBR in the hypothetical case study. This is not meant to be called directly.

Usage

```
beliaeff2002(sam_mean)
```

Arguments

`sam_mean`	a data frame, the third output of `simul`.

Value

`beliaeff2002` returns a data frame of IBR.

References

Beliaeff, B., & Burgeot, T. (2002). Integrated biomarker response: A useful tool for ecological risk assessment. Environmental Toxicology and Chemistry, 21(6), 1316–1322. doi:10.1002/etc.5620210629.

`blaise2002` *Compute Rank Sum Biomarker Index*

Description

`blaise2002` calculates RSI in the hypothetical case study. This is not meant to be called directly.

Usage

```
blaise2002(sam, sam_mean)
```

Arguments

`sam`	a data frame, the first output of `simul`.
`sam_mean`	a data frame, the third output of `simul`.

Value

`blaise2002` returns a data frame of RSI.

References

Blaise, C., Gagné, F., Pellerin, J., Hansen, P.-D., & Trottier, S. (2002). Molluscan shellfish biomarker study of the Quebec, Canada, Saguenay Fjord with the soft-shell clam, Mya arenaria. Environmental Toxicology, 17(3), 170–186. doi:10.1002/tox.10048.

`cliff` *Compute Effect Size*

Description

`cliff` calculates Cliff's delta statistic using the rank sum method.

Usage

```
cliff(v1, v0)
```

Arguments

`v1`	a vector, biomarker values from the treatment group.
`v0`	a vector, biomarker values from the control group.

Value

`cliff` returns a numeric that is the Cliff's delta of the treatment group.

References

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. Psychological Bulletin, 114(3), 494–509. doi:10.1037/00332909.114.3.494.

Vargha, A., & Delaney, H. D. (2000). A Critique and Improvement of the CL Common Language Effect Size Statistics of McGraw and Wong. Journal of Educational and Behavioral Statistics, 25(2), 101–132. doi:10.3102/10769986025002101.

Ruscio, J., & Mullen, T. (2012). Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. Multivariate Behavioral Research, 47(2), 201–223. doi:10.1080/00273171.2012.658329.

See Also

`CalcA1`.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
cliff(subset(temp$sam, Site == "S1", Bmk1, drop = TRUE),
subset(temp$sam, Site == "S0", Bmk1, drop = TRUE))
```

`compare` *Compare RSI, IBR, and Cliff's delta*

Description

`compare` calculates RSI assigned values, IBR translated scores, and Cliff's delta in the hypothetical case study.

Usage

```
compare(sam, sam_mean)
```

Arguments

`sam`	a data frame, the first output of `simul`.
`sam_mean`	a data frame, the third output of `simul`.

Value

`compare` returns a list of length 5:

`blaise`	RSI assigned values and final RSI.
`beliaeff`	IBR translated scores and final IBR.
`pham`	Cliff's delta and the average of absolute Cliff's delta.
`fig1`	ggplot object of comparisions among RSI assigned values, IBR translated scores, and Cliff's delta.
`fig2`	ggplot object of comparision among RSI, IBR, and the average of absolute Cliff's delta.

References

Blaise, C., Gagné, F., Pellerin, J., Hansen, P.-D., & Trottier, S. (2002). Molluscan shellfish biomarker study of the Quebec, Canada, Saguenay Fjord with the soft-shell clam, Mya arenaria. Environmental Toxicology, 17(3), 170–186. doi:10.1002/tox.10048.

Beliaeff, B., & Burgeot, T. (2002). Integrated biomarker response: A useful tool for ecological risk assessment. Environmental Toxicology and Chemistry, 21(6), 1316–1322. doi:10.1002/etc.5620210629.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
compare(temp$sam, temp$sam_mean)
 #might take more than 5s in some machines
```

---

ggdot *Make Dot Plot*

Description

ggdot creates dot plot of the average of absolute Cliff's delta. This is not meant to be called directly.

Usage

```
ggdot(dat, hax, vax)
```

Arguments

dat	a data frame with at least two columns.
hax	a character, name of the column to be used as the horizontal axis.
vax	a character, name of the column to be used as the vertical axis.

Value

ggdot returns a ggplot object.

---

ggheat *Make Heatmap*

Description

ggheat creates heatmaps of the Cliff's delta and S-value. This is not meant to be called directly.

Usage

```
ggheat(
  dat,
  hax,
  vax,
  cell,
  nm,
  lim,
  lo,
  hi,
  diverging = FALSE,
  env = parent.frame()
)
```

Arguments

dat	a data frame with at least three columns.
hax	a character, name of the column to be used as the horizontal axis.
vax	a character, name of the column to be used as the vertical axis.
cell	a character, name of the column to be used as the cells.
nm	a character, name of the heatmap.
lim	a numeric vector, limits of the color scale.
lo	a character, color of the color scale low end.
hi	a character, color of the color scale high end.
diverging	a logical, whether to use diverging color gradient.
env	an environment, to access outer scope variables.

Value

ggheat returns a ggplot object.

---

mbr *Compute Cliff's delta and S-value*

Description

mbr summarizes Cliff's delta and S-value for multiple groups and multiple biomarkers.

Usage

```
mbr(df)
```

Arguments

df	a data frame with the name of experimental groups or biomonitoring sites as the first column and the measurement of biomarkers as the remaining columns.

Details

The header of the first column can be any character, for example, 'group' or 'site'. The first name appearing in the first column will determine the control group or the reference site. The other names will be treatment groups or test sites. The header of the remaining columns will define the list of biomarkers.

Value

mbr returns a list of length 3:

mess	a list of length 3 confirms the information about df.
es	a data frame with 9 columns:
	test_site treatment groups or test sites.
	ref_site control group or reference site.
	t_size the sample size of treatment group or test sites.
	r_size the sample size of control group or reference site.
	biomarker individual biomarker.
	delta the Cliff's delta of treatment group or reference site.
	delta.abs the absolute Cliff's delta.
	pval the P-Value.
	sval the surprisal or S-Value.
idx	a data frame summarizes delta.abs and their average.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
mbr(temp$sam)
 #might take more than 5s in some machines
```

mbr.cliff *Compute Cliff's delta simplified*

Description

`mbr.cliff` summarizes Cliff's delta for multiple groups and multiple biomarkers.

Usage

```
mbr.cliff(df)
```

Arguments

`df`	a data frame with the name of experimental groups or biomonitoring sites as the first column and the measurement of biomarkers as the remaining columns.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
mbr.cliff(temp$sam)
 #might take more than 5s in some machines
```

mbr.glass *Compute Glass's delta simplified*

Description

`mbr.glass` summarizes Glass's delta for multiple groups and multiple biomarkers.

Usage

```
mbr.glass(df)
```

Arguments

`df`	a data frame with the name of experimental groups or biomonitoring sites as the first column and the measurement of biomarkers as the remaining columns.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
mbr.glass(temp$sam)
 #might take more than 5s in some machines
```

---

plotsam *Visualize Hypothetical Samples*

Description

plotsam plots the sample dataset of biomarker responses. This is used for the hypothetical case study.

Usage

```
plotsam(pop_mean_long, pop_profile, sam_long)
```

Arguments

pop_mean_long	a data frame, the second output of setpop.
pop_profile	a data frame, the third output of setpop.
sam_long	a data frame, the second output of simul.

Value

plotsam returns a ggplot object.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
plotsam(setting$pop_mean_long, setting$pop_profile, temp$sam_long)
```

`resampling` *Measure Statistical Uncertainty*

Description

`resampling` performs randomization test to calculate P-value and S-value.

Usage

```
resampling(v1, v0, nrand = 1999, seed = 1)
```

Arguments

`v1`	a vector, biomarker values from the treatment group.
`v0`	a vector, biomarker values from the control group.
`nrand`	an integer, the number of randomization samples. The default value is 1999.
`seed`	an integer, the seed for random number generation. Setting a seed ensures the reproducibility of the result. See `set.seed` for more details.

Value

`resampling` returns a one-row data frame with 3 numerics:

`delta`	the Cliff's delta of the treatment group.
`pval`	the observed P-value p under the null hypothesis.
`sval`	the S-value s calculated from P-value p.

References

Greenland, S. (2019). Valid P-Values Behave Exactly as They Should: Some Misleading Criticisms of P-Values and Their Resolution With S-Values. The American Statistician, 73(sup1), 106–114. doi:10.1080/00031305.2018.1529625.

Phipson, B., & Smyth, G. K. (2010). Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. Statistical Applications in Genetics and Molecular Biology, 9(1). doi:10.2202/15446115.1585.

See Also

`A1`.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
resampling(subset(temp$sam, Site == "S1", Bmk1, drop = TRUE),
subset(temp$sam, Site == "S0", Bmk1, drop = TRUE))
```

---

setpop *Define Hypothetical Populations*

Description

setpop sets the true means of biomarker responses in populations. This is used for the hypothetical case study.

Usage

```
setpop()
```

Value

setpop returns a list of length 3:

pop_mean	true means of biomarker responses in populations.
pop_mean_long	true means of biomarker responses in long format.
pop_profile	profile of biomarkers.

---

simul *Generate Hypothetical Samples*

Description

simul yields a sample dataset of biomarker responses. This is used for the hypothetical case study.

Usage

```
simul(pop_mean, size = 75)
```

Arguments

pop_mean	a data frame, the first output of setpop.
size	an integer, the sample size.

Value

simul returns a list of length 3:

sam	sample dataset.
sam_long	sample dataset in long format.
sam_mean	sample means of biomarker responses.

---

visual *Visualize Cliff's delta and S-value*

Description

visual plots Cliff's delta and S-value for multiple groups and multiple biomarkers.

Usage

```
visual(rs, rotate = FALSE, display = TRUE)
```

Arguments

rs	a list, output of mbr.
rotate	a logical, whether to rotate the biomarker labels in figures.
display	a logical, whether to display cell values in heatmaps.

Value

visual returns a list of ggplot objects:

fig.delta	heatmap of Cliff's delta.
fig.sval	heatmap of S-value.
fig.avg	dot plot of the average of absolute Cliff's delta.
mbr_fig	combined heatmaps of Cliff's delta and S-value.

Examples

```
set.seed(1)
setting <- setpop()
temp <- simul(setting$pop_mean)
mbr_result <- mbr(temp$sam)
visual(mbr_result)
 #might take more than 5s in some machines
```

Index