Package ‘boutliers’

September 21, 2025

Type Package

Title Outlier Detection and Influence Diagnostics for Meta-Analysis

Version 2.1-1

Date 2025-09-20

Maintainer Hisashi Noma <`noma@ism.ac.jp`>

Description Computational tools for outlier detection and influence diagnostics in meta-analysis (Noma et al. (2025) <doi:10.1101/2025.09.18.25336125>). Bootstrap distributions of influence statistics are computed, and explicit thresholds for identifying outliers are provided. These methods can also be applied to the analysis of influential centers or regions in multicenter or multiregional clinical trials (Aoki and Noma (2021) <doi:10.1080/24709360.2021.1921944>, Nakamura and Noma (2021) <doi:10.5691/jjb.41.117>).

Imports stats, metafor, MASS

License GPL-3

Encoding UTF-8

LazyData true

NeedsCompilation no

Author Hisashi Noma [aut, cre],
Kazushi Maruo [aut],
Masahiko Gosho [aut]

Repository CRAN

Date/Publication 2025-09-20 23:10:02 UTC

Contents

`boutliers-package` *The 'boutliers' package.*

Description

Computational tools for outlier detection and influence diagnostics in meta-analysis. Bootstrap distributions of influence statistics are computed, and explicit thresholds for identifying outliers are provided. These methods can also be applied to the analysis of influential centers or regions in multicenter or multiregional clinical trials.

References

Aoki, M., Noma, H., and Gosho, M. (2021). Methods for detecting outlying regions and influence diagnosis in multi-regional clinical trials. *Biostatistics & Epidemiology.* **5**(1): 30-48. doi:10.1080/24709360.2021.1921944

Hedges, L. V., and Olkins, I. (1985). *Statistical Methods for Meta-Analysis.* New York: Academic Press.

Nakamura, R., and Noma, H. (2021). Detection of outlying centers and influence diagnostics for the analysis of multicenter clinical trials (in Japanese). *Japanese Journal of Biometrics.* **41**(2): 117-136. doi:10.5691/jjb.41.117

Noma, H., Gosho, M., Ishii, R., Oba, K., and Furukawa, T. A. (2020). Outlier detection and influence diagnostics in network meta-analysis. *Research Synthesis Methods.* **11**(6): 891-902. doi:10.1002/jrsm.1455

Noma, H., Maruo, K., and Gosho, M. (2025). boutliers: R package of outlier detection and influence diagnostics for meta-analysis. *medRxiv.* doi:10.1101/2025.09.18.25336125

Viechtbauer, W., and Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods.* **1**(2): 112-125. doi:10.1002/jrsm.11

`finasteride` *A multicenter clinical trial data assessing the treatment effect of finasteride for benign prostatic hyperplasia*

Description

- `center`: Center ID
- `n1`: Number of observations in finasteride group
- `m1`: Mean of the change of Boyarsky score from baseline in finasteride group
- `s1`: SD of the change of Boyarsky score from baseline in finasteride group
- `n0`: Number of observations in placebo group
- `m0`: Mean of the change of Boyarsky score from baseline in placebo group
- `s0`: SD of the change of Boyarsky score from baseline in placebo group

Usage

```
data(PPI)
```

Format

A data frame with 29 rows and 7 variables

References

Nakamura, R., and Noma, H. (2021). Detection of outlying centers and influence diagnostics for the analysis of multicenter clinical trials (in Japanese). *Japanese Journal of Biometrics.* **41**(2): 117-136. doi:10.5691/jjb.41.117

Gormley, G. J., Stoner, E., Bruskewitz, R. C., et al. (1992). The effect of finasteride in men with benign prostatic hyperplasia. The Finasteride Study Group. *New England Journal of Medicine.* **327**: 1185-1191. doi:10.1056/nejm199210223271701

Gould, A. L. (1998). Multi-centre trial analysis revisited. *Statistics in Medicine.* **17**: 1779-1797.

`LRT` *Likelihood ratio test using a mean-shifted model*

Description

Implementing the likelihood ratio tests using the mean-shifted model. The bootstrap p-values are provided.

Usage

```
LRT(y, v, model="RE", data, B=2000, alpha=0.05, seed=123456)
```

Arguments

`y`	A vector of the outcome measure estimates (e.g., MD, SMD, log OR, log RR, RD)
`v`	A vector of the variance estimate of `y`
`model`	A logical value specifying the pooling model (`RE`: random-effects model, `FE`: fixed-effect model)
`data`	An optional data frame containing the variables `y` and `v`.
`B`	The number of bootstrap resampling (default: 2000)
`alpha`	The significance level (default: 0.05)
`seed`	A numeric value that determines the random seed for reproducibility (default: 123456).

Value

Results of the likelihood ratio tests involving bootstrap p-values. The outputs are ordered by the p-values.

- `id`: ID of the study.
- `LR`: The likelihood ratio statistic for based on the mean-shifted model.
- `Q`: `1-alpha`th percentile for the bootstrap distribution of the likelihood ratio statistic.
- `P`: The bootstrap p-value for the likelihood ratio statistic.

Examples

```
require(metafor)
data(SMT)

edat2 <- escalc(m1i=m1,sd1i=s1,n1i=n1,m2i=m2,sd2i=s2,n2i=n2,measure="MD",data=SMT)

LRT(yi, vi, data=edat2, B=10)
# Random-effects model.
# This is an example command for illustration. B should be >= 1000.

LRT(yi, vi, data=edat2, model="FE", B=10)
# Fixed-effect model.
# This is an example command for illustration. B should be >= 1000.
```

`PPI` *Crocker et al. (2018)'s patient and public involvement (PPI) intervention data*

Description

- `ID`: Study ID
- `d1`: Number of events in PPI intervention group
- `n1`: Number of observations in PPI intervention group
- `d2`: Number of events in non-PPI intervention group
- `n2`: Number of observations in non-PPI intervention group

Usage

```
data(PPI)
```

Format

A data frame with 21 rows and 5 variables

References

Crocker, J. C., Ricci-Cabello, I., Parker, A., Hirst, J. A., Chant, A., Petit-Zeman, S., Evans, D., Rees, S. (2018). Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. *BMJ.* **363**: k4738. doi:10.1136/bmj.k4738

`SMT` *Rubinstein et al. (2019)'s chronic low back pain data*

Description

- `ID`: Study ID
- `Souce`: First author name and year of publication
- `m1`: Estimated mean in experimental group
- `s1`: Standard deviation in experimental group
- `n1`: Number of observations in experimental group
- `m2`: Estimated mean in control group
- `s2`: Standard deviation in control group
- `n2`: Number of observations in control group

Usage

```
data(SMT)
```

Format

A data frame with 23 rows and 8 variables

References

Rubinstein, S. M,, de Zoete, A., van Middelkoop, M., Assendelft, W. J. J., de Boer, M. R., van Tulder, M. W. (2019). Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. *BMJ.* **364**: l689. doi:10.1136/bmj.l689

`STR` *Studentized residuals by leave-one-out analysis*

Description

Calculating the studentized residuals by leave-one-out analysis (studentized deleted residuals) and the percentiles of their bootstrap distributions.

Usage

```
STR(y, v, method="REML", data, B=2000, alpha=0.95, seed=123456)
```

Arguments

`y`	A vector of the outcome measure estimates (e.g., MD, SMD, log OR, log RR, RD)
`v`	A vector of the variance estimate of `y`
`method`	A logical value specifying the estimation method (default: `REML`). The same options as those available for the `method` argument of the `rma` function in the `metafor` package can be used (e.g., `FE` for the fixed-effect model, `SJ` for the Sidik–Jonkman method, and `PM` for the Paule–Mandel method).
`data`	An optional data frame containing the variables `y` and `v`.
`B`	The number of bootstrap resampling (default: 2000)
`alpha`	The bootstrap percentiles to be outputted; 0.5(1-alpha)th and (1-0.5(1-alpha))th pecentiles. Default is 0.95; 2.5th and 97.5th percentiles are calculated.
`seed`	A numeric value that determines the random seed for reproducibility (default: 123456).

Value

The studentized residuals by leave-one-out analysis. The outputs are ordered by the sizes of the studentized residuals.

- `id`: ID of the study.
- `psi`: The studentized residuals by leave-one-out analysis (studentized deleted residuals).
- `Q1`: 0.5(1-alpha)th percentile for the bootstrap distribution of the studentized residual (default: 2.5th percentile).
- `Q2`: 1-0.5(1-alpha)th percentile for the bootstrap distribution of the studentized residual (default: 97.5th percentile).

Examples

```
require(metafor)
data(PPI)

edat1 <- escalc(ai=d1,n1i=n1,ci=d2,n2i=n2,measure="OR",data=PPI)

STR(yi, vi, data=edat1, B=10)
# Random-effects model (REML estimation).
# This is an example command for illustration. B should be >= 1000.

STR(yi, vi, data=edat1, method="SJ",B=10)
# Random-effects model (Sidik-Jonkman method).
# This is an example command for illustration. B should be >= 1000.

STR(yi, vi, data=edat1, method="FE",B=10)
# Fixed-effects model.
# This is an example command for illustration. B should be >= 1000.
```

`VRATIO` *Variance ratio influential statistics*

Description

Calculating the variance ratio influential statistics by leave-one-out analysis and the percentiles of their bootstrap distributions.

Usage

```
VRATIO(y, v, method="REML", data, B=2000, alpha=0.05, seed=123456)
```

Arguments

`y`	A vector of the outcome measure estimates (e.g., MD, SMD, log OR, log RR, RD)
`v`	A vector of the variance estimate of `y`
`method`	A logical value specifying the estimation method (default: `REML`). The same options as those available for the `method` argument of the `rma` function in the `metafor` package can be used (e.g., `FE` for the fixed-effect model, `SJ` for the Sidik–Jonkman method, and `PM` for the Paule–Mandel method).
`data`	An optional data frame containing the variables `y` and `v`.
`B`	The number of bootstrap resampling (default: 2000)
`alpha`	The bootstrap percentile to be outputted (default: 0.05)
`seed`	A numeric value that determines the random seed for reproducibility (default: 123456).

Value

The variance ratio influential statistics by leave-one-out analysis and their bootstrap percentiles. The outputs are ordered by the sizes of the variance ratio statistics.

- `id`: ID of the study.
- `VR`: The VRATIO statistic (relative change of the variance of the overall estimator) by leave-one-out analysis.
- `Q1`: `alpha`th percentile for the bootstrap distribution of the VRATIO statistic.
- `TR`: The TAU2RATIO statistic (relative change of the heterogeneity variance) by leave-one-out analysis.
- `Q2`: `alpha`th percentile for the bootstrap distribution of the TAU2RATIO statistic.

Examples

```
require(metafor)
data(finasteride)

edat3 <- escalc(m1i=m1,sd1i=s1,n1i=n1,m2i=m0,sd2i=s0,n2i=n0,
measure="MD",data=finasteride)

VRATIO(yi, vi, data=edat3, B=10)
# This is an example command for illustration. B should be >= 1000.
```

Index