Package 'wasthub'

May 9, 2023

Type Package

Title Robust change-plane testing and learning based on Huber loss

Version 1.0.1

Author Xu Liu [aut,cre]

Maintainer Xu Liu <liu.xu@sufe.edu.cn>

Description Provide a method to calculate p-value of the test statistic for subgroup detecting in in the robust linear regression. In the paper Liu (2023), we propose a novel U-like statistic by taking the weighted average over the nuisance parametric space. The proposed test statistics not only improve power, but also save dramatically computational time. Many common and useful models are considered, including models with change point or change plane. We propose a novel U-like test statistic to detect multiple change planes in the robust linear regression.

License GPL (>= 2)

Depends R (>= 3.2.0)

LazyData true

NeedsCompilation yes

Repository github

URL https://github.com/xliusufe/wasthub

Encoding UTF-8

Archs i386, x64

R topics documented:

wasthub-package *Robust change-plane testing and learning based on Huber loss*

Description

Provide a method to calculate p-value of the test statistic for subgroup detecting in in the robust linear regression. In the paper Liu (2023), we propose a novel U-like statistic by taking the weighted average over the nuisance parametric space. The proposed test statistics not only improve power, but also save dramatically computational time. Many common and useful models are considered, including models with change point or change plane. We propose a novel U-like test statistic to detect multiple change planes in the robust linear regression.

Details

Package:	wasthub
Type:	Package
Version:	1.0.1
Date:	2022-05-5
License:	GPL (>= 2)

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

esthub *Estimation in Robust Linear Models with subgroups*

Description

Provide estimators of coefficients in robust linear models with subgroups.

Usage

```
esthub(x, y, method = "adaptive", maxIter = 100,
               tol = 0.00001, M = 100, tauk = NULL)
```

Arguments

x	A matrix in $R^{n\times p}$. The design matrix.
y	A vector in R^n. The response.
method	There are there methods, including 'adaptive', 'CV' and preset qq = `1.345*median( abs(y -median(y)) )/qnorm(0.75)`. Default is `method = 'adaptive'`.
maxIter	An integer, the maximum number of iterations. Default is `maxIter = 100`.
tol	Convergence threshhold. Default is `tol = 0.00001`.
M	An integer, the length of tauk. Default is `M = 100`.
tauk	A numeric vector, which is the preset τ if `method = 'CV'`. Default is `tauk = NULL`, which is `tauk = seq(0.01,10,length.out = M)`.

Details

Robust linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \epsilon_i.$$

Value

alpha	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
tau	Optimal τ.
sigma2	Estimator of the error's variance.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- esthub(x = data_gaussian$X, y = data_gaussian$Y, method = "adaptive")
fit$alpha

data(simulatedData_quantile)
fit <- esthub(x = data_quantile$X, y = data_quantile$Y, method = "adaptive")
fit$alpha
```

esthubcp *Estimation in Robust Linear Models with subgroups*

Description

Provide estimators of coefficients in robust linear models with subgroups.

Usage

```
esthubcp(data, method = "adaptive", smooth = 'sigmoid', h = NULL,
                        maxIter = 100, tol = 0.00001, M = 100, tauk = NULL)
```

Arguments

data	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
method	There are there methods, including 'adaptive', 'CV' and preset qq = `1.345*median( abs(y -median(y)) )/qnorm(0.75)`. Default is `method = 'adaptive'`.
smooth	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
h	A numeric number, which is the bandwidth in the smooth function. Default is `h = NULL`, which is `h = log(n)/sqrt(n)`.
maxIter	An integer, the maximum number of iterations. Default is `maxIter = 100`.
tol	Convergence threshhold. Default is `tol = 0.00001`.
M	An integer, the length of tauk. Default is `M = 100`.
tauk	A numeric vector, which is the preset τ if `method = 'CV'`. Default is `tauk = NULL`, which is `tauk = seq(0.01,10,length.out = M)`.

Details

Robust linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\theta} \geq 0) + \epsilon_i.$$

Value

alpha	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
beta	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
theta	Estimator of the grouping parameter $\boldsymbol{\theta}$.
delta	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
tau	Optimal τ.
sigma2	Estimator of the error's variance.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- esthubcp(data = data_gaussian, method = "adaptive")
fit$alpha

data(simulatedData_quantile)
fit <- esthubcp(data = data_quantile, method = "adaptive")
fit$beta
```

`esthubcpBoot` *Estimation in Robust Linear Models with subgroups*

Description

Provide estimators of coefficients in robust linear models with subgroups.

Usage

```
esthubcpBoot(data, method = "adaptive", smooth = 'sigmoid',
     weights = 'exponential', h = NULL, maxIter = 100, tol = 0.00001, B = 1000)
```

Arguments

`data`	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
`method`	There are there methods, including 'adaptive', 'CV' and preset qq = `1.345*median( abs(y -median(y)) )/qnorm(0.75)`. Default is `method = 'adaptive'`.
`smooth`	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
`weights`	The weights. Either "exponential" (the default), "norm", or "bernoulli", see details below.
`h`	A numeric number, which is the bandwidth in the smooth function. Default is `h = NULL`, which is `h = log(n)/sqrt(n)`.
`maxIter`	An integer, the maximum number of iterations. Default is `maxIter = 100`.
`tol`	Convergence threshhold. Default is `tol = 0.00001`.
`B`	An integer, the number of bootstrap sample sets. Default is `B = 1000`.

Details

Robust linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\theta} \geq 0) + \epsilon_i.$$

Value

`alpha`	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
`beta`	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
`theta`	Estimator of the grouping parameter $\boldsymbol{\theta}$.
`delta`	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
`tau`	Optimal τ.
`sigma2`	Estimator of the error's variance.
`std`	A vector with length $p + q + r - 1$. The standard deviation (`sd`) of parameter $(\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T, \boldsymbol{\gamma}_{-1}^T)^T$, where $\boldsymbol{\gamma}_{-1} = (\gamma_2, \cdots, \gamma_r)^T$.
`alphaB`	A matrix in $R^{p_1 \times B}$, each colum of which is the estimator of the baseline parameter $\boldsymbol{\alpha}$.
`betaB`	A matrix in $R^{p_2 \times B}$, each colum of which is the estimator of the grouping difference parameter $\boldsymbol{\beta}$.
`thetaB`	A matrix in $R^{p_3 \times B}$, each colum of which is the estimator of the grouping parameter $\boldsymbol{\theta}$.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- esthubcpBoot(data = data_gaussian)
fit$alpha

data(simulatedData_quantile)
fit <- esthubcpBoot(data = data_quantile)
fit$beta
```

esthubmcp *Estimation in Robust Linear Models with subgroups*

Description

Provide estimators of coefficients in robust linear models with subgroups.

Usage

```
esthubmcp(data, ng = 2, method = "adaptive", smooth = 'sigmoid',
          h = NULL, maxIter = 100, tol = 0.00001, M = 100, tauk = NULL)
```

Arguments

data	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
ng	An integer, which is the number of change-planes. Default is ng = 2.
method	There are there methods, including 'adaptive', 'CV' and preset qq = 1.345*median(abs(y -median(y)))/qnorm(0.75). Default is method = 'adaptive'.
smooth	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
h	A numeric number, which is the bandwidth in the smooth function. Default is h = NULL, which is h = log(n)/sqrt(n).
maxIter	An integer, the maximum number of iterations. Default is maxIter = 100.
tol	Convergence threshhold. Default is tol = 0.00001.
M	An integer, the length of tauk. Default is M = 100.
tauk	A numeric vector, which is the preset τ if method = 'CV'. Default is tauk = NULL, which is tauk = seq(0.01,10,length.out = M).

Details

Robust linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\theta} \geq 0) + \epsilon_i.$$

Value

alpha	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
beta	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
theta	Estimator of the grouping parameter $\boldsymbol{\theta}$.
delta	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
ha	Estimator of the threshholds $\{a_1, \cdots, a_S\}$, where S equals to ng.
tau	Optimal τ.
sigma2	Estimator of the error's variance.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- esthubmcp(data = data_gaussian, method = "adaptive")
fit$alpha

data(simulatedData_quantile)
fit <- esthubmcp(data = data_quantile, method = "adaptive")
fit$beta
```

`esthubmcpBoot` *Estimation in Robust Linear Models with subgroups*

Description

Provide estimators of coefficients in robust linear models with subgroups.

Usage

```
esthubmcpBoot(data, ng = 2, method = "adaptive", smooth = 'sigmoid',
     weights = 'exponential', h = NULL, maxIter = 100, tol = 0.00001, B = 1000)
```

Arguments

`data`	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
`ng`	An integer, which is the number of change-planes. Default is `ng = 2`.
`method`	There are there methods, including 'adaptive', 'CV' and preset `qq = 1.345*median( abs(y -median(y)) )/qnorm(0.75)`. Default is `method = 'adaptive'`.
`smooth`	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
`weights`	The weights. Either "exponential" (the default), "norm", or "bernoulli", see details below.
`h`	A numeric number, which is the bandwidth in the smooth function. Default is `h = NULL`, which is `h = log(n)/sqrt(n)`.
`maxIter`	An integer, the maximum number of iterations. Default is `maxIter = 100`.
`tol`	Convergence threshhold. Default is `tol = 0.00001`.
`B`	An integer, the number of bootstrap sample sets. Default is `B = 1000`.

Details

Robust linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\theta} \geq 0) + \epsilon_i.$$

Value

alpha	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
beta	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
theta	Estimator of the grouping parameter $\boldsymbol{\theta}$.
delta	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
tau	Optimal τ.
ha	Estimator of the threshholds $\{a_1, \cdots, a_S\}$, where S equals to ng.
sigma2	Estimator of the error's variance.
std	A vector with length $p + q + r - 1$. The standard deviation (sd) of parameter $(\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T, \boldsymbol{\gamma}_{-1}^T)^T$, where $\boldsymbol{\gamma}_{-1} = (\gamma_2, \cdots, \gamma_r)^T$.
alphaB	A matrix in $R^{p_1 \times B}$, each colum of which is the estimator of the baseline parameter $\boldsymbol{\alpha}$.
betaB	A matrix in $R^{p_2 \times B}$, each colum of which is the estimator of the grouping difference parameter $\boldsymbol{\beta}$.
thetaB	A matrix in $R^{p_3 \times B}$, each colum of which is the estimator of the grouping parameter $\boldsymbol{\theta}$.
haB	A matrix in $R^{S \times B}$, each colum of which is the estimator of the threshholds $\{a_1, \cdots, a_S\}$, where S equals to ng.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- esthubmcpBoot(data = data_gaussian)
fit$alpha

data(simulatedData_quantile)
fit <- esthubmcpBoot(data = data_quantile)
fit$beta
```

`estolscp` *Estimation in the Linear Models with subgroups*

Description

Provide estimators of coefficients in the linear models with subgroups.

Usage

```
estolscp(data, smooth = 'sigmoid', isBoot = FALSE, isWB = FALSE,
                    h = NULL, maxIter = 100, tol = 0.00001, B = 1000)
```

Arguments

`data`	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
`smooth`	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
`isBoot`	A bool value. A bootstrap method is used if `isBoot = TRUE`. Default is `isBoot = FALSE`.
`isWB`	A bool value. The wild bootstrap method is used if `isWB = TRUE`. Default is `isWB = FALSE`.
`h`	A numeric number, which is the bandwidth in the smooth function. Default is `h = NULL`, which is `h = log(n)/sqrt(n)`.
`maxIter`	An integer, the maximum number of iterations. Default is `maxIter = 100`.
`tol`	Convergence threshhold. Default is `tol = 0.00001`.
`B`	An integer, the number of bootstrap sample sets. Default is `B = 1000`.

Details

The linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\theta} \geq 0) + \epsilon_i.$$

Value

`alpha`	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
`beta`	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
`theta`	Estimator of the grouping parameter $\boldsymbol{\theta}$.
`delta`	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
`sigma2`	Estimator of the error's variance.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- estolscp(data = data_gaussian, smooth = 'sigmoid')
fit$alpha

data(simulatedData_quantile)
fit <- estolscp(data = data_quantile, smooth = 'sigmoid')
fit$alpha
```

---

estolsmcp *Estimation in the Linear Models with subgroups*

Description

Provide estimators of coefficients in the linear models with subgroups.

Usage

```
estolsmcp(data, ng = 2, smooth = 'sigmoid', isBoot = FALSE, isWB = FALSE,
                          h = NULL, maxIter = 100, tol = 0.00001, B = 1000)
```

Arguments

data	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
ng	An integer, which is the number of change-planes. Default is ng = 2.
smooth	The smooth function. Either "sigmoid" (the default), "pnorm", or "mixnorm", see details below.
isBoot	A bool value. A bootstrap method is used if isBoot = TRUE. Default is isBoot = FALSE.
isWB	A bool value. The wild bootstrap method is used if isWB = TRUE. Default is isWB = FALSE.
h	A numeric number, which is the bandwidth in the smooth function. Default is h = NULL, which is h = log(n)/sqrt(n).
maxIter	An integer, the maximum number of iterations. Default is maxIter = 100.
tol	Convergence threshhold. Default is tol = 0.00001.
B	An integer, the number of bootstrap sample sets. Default is B = 1000.

Details

The linear models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\sum_{s=1}^{S}\boldsymbol{\beta}_s\mathbf{1}(U_i + \boldsymbol{U}_{2i}^T\boldsymbol{\theta}_{-1} \geq a_s) + \epsilon_i.$$

with the identifiable restraint that $a_1 < a_2 < \cdots < a_S$.

Value

`alpha`	Estimator of the baseline parameter $\boldsymbol{\alpha}$.
`beta`	Estimator of the grouping difference parameter $\boldsymbol{\beta}$.
`theta`	Estimator of the grouping parameter $\boldsymbol{\theta}$.
`delta`	A vector with length n. Estimator of the indicator function $I(\boldsymbol{U}^T\boldsymbol{\theta} \geq 0)$.
`ha`	Estimator of the threshholds $\{a_1, \cdots, a_S\}$, where S equals to `ng`.
`sigma2`	Estimator of the error's variance.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
fit <- estolsmcp(data = data_gaussian)
fit$alpha

data(simulatedData_quantile)
fit <- estolsmcp(data = data_quantile)
fit$alpha
```

`exams` *Examples for Subgroup Test in Robust Linear Models*

Description

Examples for robust test of the linear regression models.

Usage

```
exams(method = "wast", B = 1000, K = 1000)
```

Arguments

method	There are there methods, including the proposed 'wast', 'sst', and 'slrt'.
B	An integer, the number of bootstrap samples.
K	An integer, the number of threshold values for 'sst' and 'slrt'.

Value

pvals	P-value of the corresponding test statistic.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2022). Change-plane testing in the robust estimating equations. Manuscript.

Examples

```
pvals <- exams(method = "wast")
pvals

pvals <- exams(method = "wastcv")
pvals
```

---

pvalhuber *P-value for Subgroup Test in Robust Linear Models*

Description

Provide p-value for subgroup test in robust linear models, including four methods 'wast', 'wastcv', 'sst', and 'slrt'.

Usage

```
pvalhuber(data, method = "wast", isWB = FALSE, B = 1000, K = 1000,
        isBeta = FALSE, shape1 = 1, shape2 = 1, N0 = 5000, MU = NULL, ZK = NULL)
```

Arguments

data	A list, including Y (response), X (baseline variable), Z (grouping difference variable), and U (grouping variable).
method	There are there methods, including the proposed 'wast', 'wastcv', 'wastapprox', 'sst', and 'slrt'.
isWB	A bool value. The wild bootstrap method is used if isWB = TRUE. Default is isWB = FALSE.
B	An integer, the number of bootstrap samples.

K	An integer, the number of threshold values for 'sst' and 'slrt' (Default is K = 1000), or the length of preset τ for 'wastcv' (Default is K = 100).
isBeta	A bool value. The weight $w(\gamma)$ is chosen to be Beta distribution if isBeta = TRUE, which can be used if the grouping difference variable is bounded in $[0, 1]$. Default is FALSE.
shape1	The first parameter of Best distribution if isBeta = TRUE.
shape2	The second parameter of Best distribution if isBeta = TRUE.
N0	An integer, the number of samples to approximate ω_{ij} for 'wastapprox'. Default is N0 = 5000.
MU	A vector with same length as $\boldsymbol{U}$, which is the mean of weight to approximate ω_{ij} for 'wastapprox'. Default is MU = NULL, in which MU = runif(p3) -0.5.
ZK	A vector with length N0, which is normal sample to approximate ω_{ij} for 'wastapprox'. Default is ZK = NULL, in which ZK = rnorm(N0).

Details

Generalized linear models

$$y_i = \boldsymbol{X}_i^T \boldsymbol{\alpha} + \boldsymbol{Z}_i^T \boldsymbol{\beta} \mathbf{1}(\boldsymbol{U}_i^T \boldsymbol{\gamma} \geq 0) + \epsilon_i.$$

The hypothesis test problem is

$$H_0 : \boldsymbol{\beta} = \mathbf{0} \quad versus \quad H_1 : \boldsymbol{\beta} \neq \mathbf{0}.$$

Value

pvals	P-value of the corresponding test statistic.

References

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62(6):1383-1414.

Fan, A., Rui, S., and Lu, W. (2017). Change-plane analysis for subgroup detection and sample size calculation. Journal of the American Statistical Association, 112(518):769-778.

Huang, Y., Cho, J., and Fong, Y. (2021). Threshold-based subgroup testing in logistic regression models in two phase sampling designs. Journal of the Royal Statistical Society: Series C. 291-311.

Liu, X. (2023). Change-plane testing in the generalized estimating equations. Manuscript.

Liu. X. (2023). Robust change-plane testing and learning based on Huber loss. Manuscript.

Examples

```
data(simulatedData_gaussian)
pvals <- pvalhuber(data = data_gaussian, method = "wast")
pvals

data(simulatedData_quantile)
pvals <- pvalhuber(data = data_quantile, method = "wast")
pvals
```

simulatedData *Simulated data from robust linear regression models*

Description

Simulated data from the robust linear regression, including simulatedData_gaussian and simulatedData_quantile.

Usage

```
data(simulatedData_gaussian)
```

Details

We simulated data robust linear regression models

$$y_i = \boldsymbol{X}_i^T\boldsymbol{\alpha} + \boldsymbol{Z}_i^T\boldsymbol{\beta}\mathbf{1}(\boldsymbol{U}_i^T\boldsymbol{\gamma} \geq 0) + \epsilon_i.$$

- Y: the response, an n-vector
- X: the baseline variable with dimension $n \times p$
- Z: the grouping difference variable with dimension $n \times q$
- U: the grouping variable with dimension $n \times r$

References

Liu, X. (2022). Change-plane testing in the generalized estimating equations. Manuscript.

Examples

```
data(simulatedData_gaussian)

y <- data_gaussian$Y[1:5]
x <- dim(data_gaussian$X)
z <- dim(data_gaussian$Z)
u <- dim(data_gaussian$U)

data(simulatedData_quantile)
y <- data_quantile$Y[1:5]
x <- dim(data_quantile$X)
z <- dim(data_quantile$Z)
u <- dim(data_quantile$U)
```

Index