Package 'TEEReg'

June 14, 2016

Type Package

Title Trimmed Elemental Estimation for Linear Models

Version 1.1

Date 2016-06-12

Author Wei Jiang and Matthew S. Mayo

Maintainer Wei Jiang <wjiang@kumc.edu>

Description For fitting multiple linear regressions, the ordinary least squares approach is sensitive to outliers and/or violations of model assumptions. The trimmed elemental estimators are more robust to such situations. This package contains functions for computing the trimmed elemental estimates, as well as for creating the bias-corrected and accelerated bootstrap confidence intervals based on elemental regressions.

License GPL (>= 2)

Imports stats, utils

NeedsCompilation no

Repository CRAN

Date/Publication 2016-06-14 08:29:46

R topics documented:

TEEReg-package *Trimmed Elemental Estimation for Linear Models.*

Description

Package provides functions for computing the trimmed elemental estimates, as well as for creating the bias-corrected and accelerated bootstrap confidence intervals based on trimmed elemental regressions. This approach offers a robust alternative to ordinary least squares.

Details

Package:	TEEReg
Type:	Package
Version:	1.1
Date:	2016-06-12
License:	GPL (version 2 or later)

There are two major functions in the package. The TEE function is for computing the trimmed elemental estimates. The TEE.BCa function is for creating the bias-corrected accelerated bootstrap confidence intervals for regression parameters.

Author(s)

Wei Jiang and Matthew S. Mayo

Maintainer: Wei Jiang <wjiang@kumc.edu>

References

M. S. Mayo and B. Gray. Elemental subsets: the building blocks of regression. The American Statistician, 51: 122-129, 1997.

M. Hall and M. S. Mayo. Bootstrap confidence intervals and coverage probabilities of regression parameter estimates using trimmed elemental estimation. Journal of Modern Applied Statistical Methods, 7: 514-525, 2008.

Examples

```
data(telephone)
fit <- TEE(formula=Y~X,data=telephone,p.trimmed=0.5,p.subsample=0.5,method="tee")
TEE.BCa(formula=Y~X,data=telephone,est.TEE=fit$coefficients,p.trimmed=0.5,p.subsample=0.5,
method="tee",conf.level=0.05,n.boot=20)
```

TEE *Compute the trimmed elemental estimates.*

Description

This function computes the trimmed elemental estimates for regression parameters.

Usage

```
TEE(formula,data,offset=NULL,p.trimmed=NULL,p.subsample=1,method="tee")
```

Arguments

formula	define a symbolic description of the model to be fitted.
data	specify the dataset used for performing regression analysis. It must be formatted in data frame.
offset	specify an known component to be included in the linear predictor during fitting. This argument should be either NULL or a numeric vector with length equal to the number of observations.
p.trimmed	define the trimming proportion of elemental subsets. This should be either NULL or a numeric value between 0 and 1. When method = "tee" is specified, a numberic value must be provided.
p.subsample	this is the proportion of subsampling with values between 0 and 1. The default value is 1, meaning that calculations are based on the full data.
method	two options are supported: "ols" stands for ordinary least squares and "tee" stands for trimmed elemental estimation.

Details

For more details about trimmed elemental estimation, see Mayo and Gray (1997).

Value

call	call to the function.
coefficients	estimated regression coefficients with intercept.
residuals	residuals resulted from the fitted model, i.e. y - yhat.
fitted.values	values fitted by the model, i.e. yhat = x betahat.

Author(s)

Wei Jiang and Matthew S. Mayo

References

M. S. Mayo and B. Gray. Elemental subsets: the building blocks of regression. The American Statistician, 51: 122-129, 1997.

Examples

```
data(telephone)
fit<-TEE(formula=Y~X,data=telephone,p.trimmed=0.5,p.subsample=0.5,method="tee")
```

`TEE.BCa` *Compute the bias-corrected accelerated bootstrap confidence intervals.*

Description

This function create the bias-corrected accelerated bootstrap confidence intervals based on the trimmed elemental regressions.

Usage

```
TEE.BCa(formula,data,offset=NULL,p.trimmed=NULL,p.subsample=1,method ="tee",
est.TEE,conf.level,n.boot)
```

Arguments

`formula`	define a symbolic description of the model to be fitted.
`data`	specify the dataset used for performing regression analysis. It must be formatted in data frame.
`offset`	specify a known component to be included in the linear predictor during fitting. This argument should be either NULL or a numeric vector with length equal to the number of observations.
`p.trimmed`	define the trimming proportion of elemental subsets. This should be either NULL or a numeric value between 0 and 1. When method = "tee" is specified, a numeric value must be provided.
`p.subsample`	this is the proportion of subsampling with values between 0 and 1. The default value is 1, meaning that calculations are based on the full data.
`method`	two options are supported: "ols" stands for ordinary least squares and "tee" stands for trimmed elemental estimation.
`est.TEE`	this is for trimmed elemental regression estimates.
`conf.level`	the confidence level.
`n.boot`	number of bootstrap samples.

Details

For discussions about bootstrap confidence intervals and coverage probabilities of regression parameter estimates using trimmed elemental estimation, see Hall and Mayo (2008).

Value

call	call to the function.
ci	bias-corrected accelerated bootstrap confidence interval estimates for the regression parameters.

Author(s)

Wei Jiang and Matthew S. Mayo

References

M. Hall and M. S. Mayo. Bootstrap confidence intervals and coverage probabilities of regression parameter estimates using trimmed elemental estimation. Journal of Modern Applied Statistical Methods, 7: 514-525, 2008.

Examples

```
data(telephone)
fit <- TEE(formula=Y~X,data=telephone,p.trimmed=0.5,p.subsample=0.5,method="tee")
TEE.BCa(formula=Y~X,data=telephone,est.TEE=fit$coefficients,p.trimmed=0.5,p.subsample=0.5,
method="tee",conf.level=0.05,n.boot=20)
```

---

telephone *An example dataset for TEEReg.*

Description

The telephone data records the number of telephone calls (tens of millions) made in Belgium from 1959 to 1973. There are several outliers resulted from mistakes in recording units over the years 1964 to 1969.

Usage

```
data(telephone)
```

Format

A data frame with 24 observations on the following 2 variables.

`Y` a numeric vector: the number of telephone calls in tens of millions.

`X` a numeric vector: years.

Source

R. J. Rousseeuw and A. M. Lerory. Robust Regression and Outlier Detection. John Wiley and Sons, New York, 1987.

Examples

```
data(telephone)
plot(telephone)
```

Index