Package 'perccal'

October 14, 2022

Type Package

Title Implementing Double Bootstrap Linear Regression Confidence Intervals Using the 'perc-cal' Method

Version 1.0

Date 2016-06-14

Maintainer Daniel McCarthy <danielmc@wharton.upenn.edu>

URL http://www.danielminhmccarthy.com

Description Contains functions which allow the user to compute confidence intervals quickly using the double bootstrap-based percentile calibrated ('perc-cal') method for linear regression coefficients. 'perccal_interval()' is the primary user-facing function within this package.

License GPL-3

Imports Rcpp (>= 0.11.5)

LinkingTo Rcpp, RcppArmadillo, RcppEigen

NeedsCompilation yes

Author Daniel McCarthy [aut, cre, ctb]

Repository CRAN

Date/Publication 2016-06-14 20:34:32

R topics documented:

perccal-package *Computing Confidence Intervals Via Double Bootstrap 'perc-cal' Method*

Description

Contains functions which allow the user to compute confidence intervals quickly using the double bootstrap-based percentile calibrated ('perc-cal') method for linear regression coefficients.

Details

Package:	perccal
Type:	Package
Version:	1.0
Date:	2016-06-14
License:	GPL-3

Contains functions which allow users to compute confidence intervals quickly using the double bootstrap-based percentile calibrated ('perc-cal') method for linear regression coefficients.

The help of Justin Bleich is strongly acknowledged.

Author(s)

Daniel McCarthy

Maintainer: Daniel McCarthy <danielmc@wharton.upenn.edu>

References

Efron, Bradley; Tibshirani, Robert J. "An Introduction to the Bootstrap" 1994. Book. Publisher: CRC Press.

McCarthy, Daniel; Zhang, Kai; Berk, Richard; Brown, Lawrence; Buja, Andreas; George, Edward; Zhao, Linda. "Calibrated Percentile Double Bootstrap For Robust Linear Regression Inference" 2016. Available on arXiv: https://arxiv.org/abs/1511.00273

Cdboot_multi *Fast computation of internal double bootstrap calculations*

Description

This is the workhorse function of the package, speeding up computations within double bootstrap routine.

Usage

```
Cdboot_multi(xxyy, lgridlo, lgridhi, B, B2, G)
```

Arguments

xxyy	(n by p+1) matrix for X (design matrix) and response vector y.
lgridlo	Lower quantile values of double bootstrap distribution to obtain.
lgridhi	Upper quantile values of double bootstrap distribution to obtain.
B	Number of 1st stage bootstrap samples.
B2	Number of 2nd stage double bootstrap samples.
G	Calculate quantile-based empirical coverage at this many grid points

Value

theta_hat_boot first-level bootstrap estimates of all slope coefficients

theta_qtl_lgrid_lo (p+1 by B by G by 1) matrix for lower quantiles at all grid points for all predictors over all bootstrap samples.

theta_qtl_lgrid_hi (p+1 by B by G by 1) matrix for upper quantiles at all grid points for all predictors over all bootstrap samples.

Cquantile *Fast computation of quantiles*

Description

Helper function which takes as input a vector and obtains quantiles for it. Number of quantiles may be greater than one.

Usage

```
Cquantile(xx, p)
```

Arguments

xx	Numeric vector we are obtaining quantiles for.
p	Numeric vector of quantiles.

Value

Numeric vector containing quantiles, possibly greater than one.

perccal_interval *Calculates Percentile-Calibrated Linear Regression Confidence Intervals*

Description

This is the main function of the package. It takes as inputs the predictor/response matrix appended together, which can be either a data frame or a matrix, along with the desired coverage and other settings, and outputs marginal confidence intervals for each of the predictors, including the intercept.

Usage

```
perccal_interval(Xy, alpha, G = 20, B = 999, B2 = 999)
```

Arguments

Xy	[n by (p+1)] matrix: X in columns 1 to p, y in column p+1. X is the design matrix, and is assumed to not include a vector of one's.
alpha	Target coverage desired.
G	Number of grid points to evaluate calibrated percentile method on each side over.
B	Number of 1st stage bootstrap samples.
B2	Number of 2nd stage double bootstrap samples.

Value

Return a (p+1)x2 matrix containing confidence intervals for all regression coefficients, estimated via the perc-cal method.

Examples

```
set.seed(1234)
n = 32
B = 500
B2 = 500
G=20
x1=rnorm(n)
x2=rnorm(n)
eps=rnorm(n)
y = x1 + 2*x2 + eps
Xy = cbind(x1,x2,y)
alpha = .025
perccal_interval(Xy, alpha, G, B, B2)
```

`sample_rcpp` *Sample from [1,2,...,N] with replacement nsamp times in Rcpp.*

Description

Helper function which samples from [1,2,...,N] with replacement nsamp times in Rcpp.

Usage

```
sample_rcpp(N, nsamp)
```

Arguments

`N`	Largest integer to sample from.
`nsamp`	number of samples from [1,2,...,N] with replacement to obtain.

Value

samps nsamp-length vector of samples from [1,2,...,N] with replacement to obtain.

Index