

Package ‘hcci’

January 21, 2025

Type Package

Title Interval Estimation of Linear Models with Heteroskedasticity

Version 1.2.0

Maintainer Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

Description Calculates the interval estimates for the parameters of linear models with heteroscedastic regression using bootstrap - (Wild Bootstrap) and double bootstrap-t (Wild Bootstrap). It is also possible to calculate confidence intervals using the percentile bootstrap and percentile bootstrap double. The package can calculate consistent estimates of the covariance matrix of the parameters of linear regression models with heteroscedasticity of unknown form. The package also provides a function to consistently calculate the covariance matrix of the parameters of linear models with heteroscedasticity of unknown form. The bootstrap methods exported by the package are based on the master's thesis of the first author, available at <https://raw.githubusercontent.com/prdm0/hcci/master/references/dissertacao_mestrado.pdf>. The hcci package in previous versions was cited in the book VINOD, Hrishikesh D. Hands-on Intermediate Econometrics Using R: Templates for Learning Quantitative Methods and R Software. 2022, p. 441, ISBN 978-981-125-617-2 (hardcover). The simple bootstrap schemes are based on the works of Cribari-Neto F and Lima M. G. (2009) <[doi:10.1080/00949650801935327](https://doi.org/10.1080/00949650801935327)>, while the double bootstrap schemes for the parameters that index the linear models with heteroscedasticity of unknown form are based on the works of Beran (1987) <[doi:10.2307/2336685](https://doi.org/10.2307/2336685)>. The use of bootstrap for the calculation of interval estimates in regression models with heteroscedasticity of unknown form from a weighting of the residuals was proposed by Wu (1986) <[doi:10.1214/aos/1176350142](https://doi.org/10.1214/aos/1176350142)>. This bootstrap scheme is known as weighted or wild bootstrap.

Author Pedro Rafael Diniz Marinho [aut, cre],
Francisco Cribari Neto [aut, ctb]

License GPL (>= 3)

URL <https://github.com/prdm0/hcci>

BugReports <https://github.com/prdm0/hcci/issues>

NeedsCompilation no

Repository CRAN

Date/Publication 2025-01-21 18:50:02 UTC

Contents

hcci-package	2
HC	3
Pboot	4
QT	6
schools	7
Tboot	8
Index	10

 hcci-package

Interval Estimation of Linear Models with Heteroskedasticity

Description

Calculates the interval estimates for the parameters of linear models with heteroscedastic regression using bootstrap - (Wild Bootstrap) and double bootstrap-t (Wild Bootstrap). It is also possible to calculate confidence intervals using the percentile bootstrap and percentile bootstrap double. The package can calculate consistent estimates of the covariance matrix of the parameters of linear regression models with heteroscedasticity of unknown form. The package also provides a function to consistently calculate the covariance matrix of the parameters of linear models with heteroscedasticity of unknown form. The bootstrap methods exported by the package are based on the master's thesis of the first author, available at https://raw.githubusercontent.com/prdm0/hcci/master/references/dissertacao_mestrado.pdf. The hcci package in previous versions was cited in the book VINOD, Hrishikesh D. Hands-on Intermediate Econometrics Using R: Templates for Learning Quantitative Methods and R Software. 2022, p. 441, ISBN 978-981-125-617-2 (hardcover). The simple bootstrap schemes are based on the works of Cribari-Neto F and Lima M. G. (2009) <https://doi.org/10.1080/00949650801935327>, while the double bootstrap schemes for the parameters that index the linear models with heteroscedasticity of unknown form are based on the works of Beran (1987) <https://doi.org/10.2307/2336685>. The use of bootstrap for the calculation of interval estimates in regression models with heteroscedasticity of unknown form from a weighting of the residuals was proposed by Wu (1986) <https://dx.doi.org/10.1214/aos/1176350>. This bootstrap scheme is known as weighted or wild bootstrap.

Details

Package: hcci
 Type: Package
 Version: 1.1.0
 Date: 2024-03-30
 License: GPL (>=3)

Author(s)

Maintainer: Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

Contributor: Francisco Cribari Neto <cribari@de.ufpe.br>

References

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics and Data Analysis*, 45, 215-233.

Cribari-Neto, F.; Lima, M.G. (2009). Heteroskedasticity-consistent interval estimators. *Journal of Statistical Computation and Simulation*, 79, 787-803;

Cribari-Neto, F.; Souza, T.C.; Vasconcellos, K.L.P. (2007). Inference under heteroskedasticity and leveraged data. *Communications in Statistics, Theory and Methods*, 36, 1877-1888. [Errata: 37, 2008, 3329-3330.]

Horn, S.D.; Horn, R.A.; Duncan, D.B. (1975). Estimating heteroskedastic variances in linear models. *Journal of the American Statistical Association*, 70, 380-385.

MacKinnon, J.G.; White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite-sample properties. *Journal of Econometrics*, 29, 305-325.

McCullough, B.D; Vinod, H.D. (1998). Implementing the double bootstrap, 12, 79-95.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817-838.

Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis, 14, 1261-1295;

 HC

Covariance Matrix - (HC0, HC2, HC3, HC4 and HC5)

Description

This function calculates the covariance structure for heteroskedasticity linear regression model.

Usage

```
HC(model, method=4, k=0.7)
```

Arguments

model	Any object of class <code>lm</code> ;
method	Method HC that will be used to estimate the covariance structure. The argument method may be 0, 2, 3, 4 or 5;
k	Constant used by the method HC5. The suggestion of the authors is to use $k = 0.7$.

Value

Returns an object of class `matrix` with the estimated covariance matrix.

Author(s)

Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

References

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics and Data Analysis*, 45, 215-233.

Cribari-Neto, F.; Souza, T.C.; Vasconcellos, K.L.P. (2007). Inference under heteroskedasticity and leveraged data. *Communications in Statistics, Theory and Methods*, 36, 1877-1888. [Errata: 37, 2008, 3329-3330.]

Horn, S.D.; Horn, R.A.; Duncan, D.B. (1975). Estimating heteroskedastic variances in linear models. *Journal of the American Statistical Association*, 70, 380-385.

MacKinnon, J.G.; White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite-sample properties. *Journal of Econometrics*, 29, 305-325.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817-838.

Examples

```
data(schools)
datas = schools[-50L,]
y = datas$Expenditure
x = datas$Income/10000
model = lm(y ~ x)
HC(model, method=4)
```

Pboot

Percentile Bootstrap Confidence Interval (Wild Bootstrap) - Linear Models Heteroskedasticity

Description

This function calculates confidence intervals for the parameters in heteroskedasticity linear regression models. The intervals are estimated by bootstrap percentile.

Usage

```
Pboot(model, significance=0.05, double=FALSE, J=NULL, K=NULL,
      distribution="rademacher")
```

Arguments

model	Any object of class <code>lm</code> ;
significance	Significance level of the test. By default, the level of significance is 0.05;
double	If <code>double = TRUE</code> will be calculated intervals bootstrap t and double bootstrap t. The default is <code>double = FALSE</code> ;

J	Number of replicas of the first bootstrap;
K	Number of replicas of the second bootstrap;
distribution	Distribution of the random variable with mean zero and variance one. This random variable multiplies the error estimates in the generation of the samples. The argument <code>distribution</code> can be <code>rademacher</code> or <code>normal</code> (standard normal). The default is <code>distribution = rademacher</code> .

Value

A list with the following components:

<code>beta</code>	A numeric vector of length 2 containing the estimated coefficients of the model.
<code>ci_lower_simple</code>	A numeric vector of length 2 containing the lower bounds of the simple bootstrap confidence intervals for the coefficients.
<code>ci_upper_simple</code>	A numeric vector of length 2 containing the upper bounds of the simple bootstrap confidence intervals for the coefficients.
<code>ci_lower_double</code>	A logical vector of length 0 or 2. If <code>'double = FALSE'</code> , this will be a logical vector of length 0. If <code>'double = TRUE'</code> , this will be a numeric vector containing the lower bounds of the double bootstrap confidence intervals for the coefficients.
<code>ci_upper_double</code>	A logical vector of length 0 or 2. If <code>'double = FALSE'</code> , this will be a logical vector of length 0. If <code>'double = TRUE'</code> , this will be a numeric vector containing the upper bounds of the double bootstrap confidence intervals for the coefficients.
J	A numeric value indicating the number of bootstrap resamples used in the simple bootstrap.
K	A numeric value indicating the number of bootstrap resamples used in the double bootstrap, if <code>'double = TRUE'</code> .

Author(s)

Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

References

- Booth, J.G. and Hall, P. (1994). Monte Carlo approximation and the iterated bootstrap. *Biometrika*, 81, 331-340.
- Cribari-Neto, F.; Lima, M.G. (2009). Heteroskedasticity-consistent interval estimators. *Journal of Statistical Computation and Simulation*, 79, 787-803;
- Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis, 14, 1261-1295;
- McCullough, B.D; Vinod, H.D. (1998). Implementing the double bootstrap, 12, 79-95.

See Also

[Tboot](#).

Examples

```
data(schools)
datas = schools[-50,]
y = datas$Expenditure
x = datas$Income/10000
model = lm(y ~ x)
Pboot(model=model, significance = 0.05, double = FALSE,
       J=1000, K = 100, distribution = "rademacher")
```

QT

Quasi-t test in linear regression models

Description

This function performs the quasi-t test for the parameters that index linear regression models, considering models with unknown heteroscedasticity, where HC methods are used to estimate the covariance matrix.

Usage

```
QT(model, significance = 0.05, hc=4, h0=0, ...)
```

Arguments

model	Any object of class <code>lm</code> ;
significance	Significance level of the test. By default, the level of significance is <code>0.05</code> ;
hc	Method HC that will be used to estimate the covariance structure. The argument method may be <code>0</code> , <code>2</code> , <code>3</code> , <code>4</code> or <code>5</code> ;
h0	Constant used in the null hypothesis (default is <code>h0 = 0</code>);
...	Additional arguments to be passed to the function HC.

Value

Returns an object of class `list` with the estimated covariance matrix.

Author(s)

Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

References

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics and Data Analysis*, 45, 215-233.

Examples

```
library(hcci)
data(schools)
y = schools$Expenditure # dependent variable
x = schools$Income/10000 # regressor scaled by 10^4
model_1 = lm(y ~ x)
model_2 = lm(y ~ x+I(x^2))
QT(model_1, significance = 0.05, hc=4, h0=0)
QT(model_2, significance = 0.05, hc=4, h0=0)
```

schools

US Expenditures for Public Schools

Description

Per capita expenditure on public schools and per capita income by state in 1979.

Usage

```
data(schools)
```

Format

A data frame containing 51 observations of 2 variables.

Expenditure per capita expenditure on public schools,

Income per capita income.

References

Cribari-Neto F. (2004), Asymptotic Inference Under Heteroskedasticity of Unknown Form, *Computational Statistics and Data Analysis*, 45, 215-233.

Greene W.H. (1993), *Econometric Analysis*, 3rd edition. Macmillan Publishing Company, New York.

US Department of Commerce (1979), *Statistical Abstract of the United States*. US Government Printing Office, Washington, DC.

Tboot	<i>Bootstrap-t Confidence Interval (Wild Bootstrap) - Linear Models Heteroskedasticity</i>
-------	--

Description

This function calculates confidence intervals for the parameters in heteroskedasticity linear regression models. Ranges are estimated by the bootstrap-t and double bootstrap-t.

Usage

```
Tboot(model, significance=0.05, hc=4, double=FALSE, J=NULL, K=NULL,
      distribution="rademacher")
```

Arguments

model	Any object of class lm;
significance	Significance level of the test. By default, the level of significance is 0.05;
hc	Method HC that will be used to estimate the covariance structure. The argument method may be 0, 2, 3, 4 or 5;
double	If double = TRUE will be calculated intervals bootstrap-t and double bootstrap-t. The default is double = FALSE;
J	Number of replicas of the first bootstrap;
K	Number of replicas of the second bootstrap;
distribution	Distribution of the random variable with mean zero and variance one. This random variable multiplies the error estimates in the generation of the samples. The argument distribution can be rademacher or normal (standard normal). The default is distribution = rademacher.

Value

A list with the following components:

beta	A numeric vector of length 2 containing the estimated coefficients of the model.
ci_lower_simple	A numeric vector of length 2 containing the lower bounds of the simple bootstrap confidence intervals for the coefficients.
ci_upper_simple	A numeric vector of length 2 containing the upper bounds of the simple bootstrap confidence intervals for the coefficients.
ci_lower_double	A logical vector of length 0 or 2. If 'double = FALSE', this will be a logical vector of length 0. If 'double = TRUE', this will be a numeric vector containing the lower bounds of the double bootstrap confidence intervals for the coefficients.

ci_upper_double	A logical vector of length 0 or 2. If 'double = FALSE', this will be a logical vector of length 0. If 'double = TRUE', this will be a numeric vector containing the upper bounds of the double bootstrap confidence intervals for the coefficients.
J	A numeric value indicating the number of bootstrap resamples used in the simple bootstrap.
K	A numeric value indicating the number of bootstrap resamples used in the double bootstrap, if 'double = TRUE'.

Author(s)

Pedro Rafael Diniz Marinho <pedro.rafael.marinho@gmail.com>

References

- Booth, J.G. and Hall, P. (1994). Monte Carlo approximation and the iterated bootstrap. *Biometrika*, 81, 331-340.
- Cribari-Neto, F.; Lima, M.G. (2009). Heteroskedasticity-consistent interval estimators. *Journal of Statistical Computation and Simulation*, 79, 787-803;
- Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis, 14, 1261-1295;
- McCullough, B.D; Vinod, H.D. (1998). Implementing the double bootstrap, 12, 79-95.

See Also

[Pboot](#).

Examples

```
data(schools)
datas = schools[-50,]
y = datas$Expenditure
x = datas$Income/10000
model = lm(y ~ x)
Tboot(model=model, significance = 0.05, hc = 4, double = FALSE,
      J=1000, K = 100, distribution = "rademacher")
```

Index

* **bootstrap**

Pboot, 4

Tboot, 8

* **covariance**

HC, 3

QT, 6

* **datasets**

schools, 7

* **heteroscedasticity**

Pboot, 4

Tboot, 8

* **heteroskedasticity**

HC, 3

QT, 6

HC, 3

hcci (hcci-package), 2

hcci-package, 2

Pboot, 4, 9

QT, 6

schools, 7

Tboot, 6, 8