

Package ‘pretest’

July 23, 2025

Title A Novel Approach to Predictive Accuracy Testing in Nested Environments

Version 0.2

Description This repository contains the codes for using the predictive accuracy comparison tests developed in Pitarakis, J. (2023) <[doi:10.1017/S0266466623000154](https://doi.org/10.1017/S0266466623000154)>.

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Author Rong Peng [aut, cre]

Maintainer Rong Peng <r.peng@soton.ac.uk>

Repository CRAN

Date/Publication 2023-10-19 07:00:02 UTC

Contents

dm_cw	2
lr_var	2
Nested_Stats_S0	3
Nested_Stats_Sbar	4
recursive_hstep_fast	6
recursive_hstep_slow	7
Index	8

dm_cw

*Diebold-Mariano Test and Clark & West Test***Description**

It calculates the original DM statistics and the CW adjusted version of DM statistics, including the versions based on a Newey-West type estimator of the long run variance.

Usage

```
dm_cw(Ehat1, Ehat2)
```

Arguments

Ehat1	Residual series from Model 1 (the smaller model). One dimension and numeric.
Ehat2	Residual series from Model 2 (the larger/nested model). One dimension and numeric.

Value

A list of statistics and corresponding P values will be produced.

References

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1), 291-311.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253-263.

Examples

```
e1<- rnorm(15);
e2<- rnorm(15);
temp1 <- dm_cw(e1,e2)
```

lr_var

*Heteroskedastic Long run variance***Description**

Long-run covariance estimation using Newey-West (Bartlett) weights

Usage

```
lr_var(u, nlag = NULL, demean = TRUE)
```

Arguments

<code>u</code>	P by K vector of residual series, for which we recommend to use the recursive residuals from larger model.
<code>nlag</code>	Non-negative integer containing the lag length to use. If empty or not included, $nlag = \min(\text{floor}(1.2 \cdot T^{1/3}), T)$ will be used.
<code>demean</code>	Logical true or false (0 or 1) indicating whether the mean should be subtracted when computing.

Details

Copyright: Kevin Sheppard Kevin.sheppard@economics.ox.ac.uk Revision: 3 Date: 5/1/2007

Value

K by K vector of Long run variance using Newey-West (Bartlett) weights.

Examples

```
x<- rnorm(15);
#Newey-West covariance with automatic BW selection
lrcov = lr_var(x)
#Newey-West covariance with 10 lags
lrcov = lr_var(x, 10)
#Newey-West covariance with 10 lags and no demeaning
lrcov = lr_var(x, 10, 0)
```

Nested_Stats_S0

Predictive Accuracy Testing for Nested Environment S^0

Description

It calculates the S^0 statistics for nested models with null hypothesis being the two models having equal predictive power following Pitarakis (2023). There are in total four versions of S^0 , based on the assumptions of variance (homo or hete) and residuals (original or adjusted). All S^0 statistics will be standardised to a standard $N(0,1)$ normal distribution, and corresponding P values would be provided.

Usage

```
Nested_Stats_S0(Ehat1, Ehat2, lam10, lam20)
```

Arguments

<code>Ehat1</code>	Residual series from Model 1 (the smaller model). One dimension and numeric.
<code>Ehat2</code>	Residual series from Model 2 (the larger/nested model). One dimension and numeric.
<code>lam10</code>	Fraction of the sample used for Model 1, which should be within 0 and 1.
<code>lam20</code>	Fraction of the sample used for Model 2, which should be within 0 and 1. Note that lam10 cannot equal to lam20 (c.f. Pitarakis, 2023).

Value

A list of S^0 statistics and corresponding P values will be produced. "adj" means a Clark and West's (2007) reformulation of sample MSE has been applied, and "NW" means robust Newey-West type estimator (c.f. Deng and Perron, 2008) for heteroskedastic errors has been used.

Author(s)

Rong Peng, <r.peng@soton.ac.uk>

References

- Pitarakis, J. Y. (2023). A novel approach to predictive accuracy testing in nested environments. *Econometric Theory*, 1-44.
- Deng, A., & Perron, P. (2008). The limit distribution of the CUSUM of squares test under general mixing conditions. *Econometric Theory*, 24(3), 809-822.
- Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1), 291-311.

See Also

[Nested_Stats_Sbar](#)

Examples

```
e1<- rnorm(15);
e2<- rnorm(15);
temp1 <- Nested_Stats_S0(e1,e2,lam10=0.5,lam20=0.8)
temp1$S_lam10_lam20_adj_NW      #S^0_T(lam10, lam^20) with CW adjustment and NW correction
temp1$pv_S_lam10_lam20_adj_NW  #P value of it
```

Nested_Stats_Sbar

Predictive Accuracy Testing for Nested Environment SBar

Description

It calculates the SBar statistics for nested models with null hypothesis being the two models having equal predictive power following Pitarakis (2023). There are in total four versions of SBar, based on the assumptions of variance (homo or hete) and residuals (original or adjusted). All SBar statistics will be standardised to a standard $N(0,1)$ normal distribution, and corresponding P values would be provided.

Usage

```
Nested_Stats_Sbar(Ehat1, Ehat2, lam20, tau0)
```

Arguments

Ehat1	Residual series from Model 1 (the smaller model). One dimension and numeric.
Ehat2	Residual series from Model 2 (the larger/nested model). One dimension and numeric.
lam20	Fraction of the sample used for Model 2, which should be within 0 and 1.
tau0	Fraction to determine the user-chosen range of lam10 over which the average is taken.

Value

A list of SBar statistics and corresponding P values will be produced. "adj" means a Clark and West's (2007) reformulation of sample MSE has been applied, and "NW" means robust Newey-West type estimator (c.f. Deng and Perron, 2008) for heteroskedastic errors has been used.

Author(s)

Rong Peng, <r.peng@soton.ac.uk>

References

- Pitarakis, J. Y. (2023). A novel approach to predictive accuracy testing in nested environments. *Econometric Theory*, 1-44.
- Deng, A., & Perron, P. (2008). The limit distribution of the CUSUM of squares test under general mixing conditions. *Econometric Theory*, 24(3), 809-822.
- Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1), 291-311.

See Also

[Nested_Stats_S0](#)

Examples

```
e1<- rnorm(15);
e2<- rnorm(15);
temp1 <- Nested_Stats_S0(e1,e2, lam10=0.5, lam20=0.8)
temp1$S_lam10_lam20_adj_NW    #S^0_T(lam10, lam^20) with CW adjustment and NW correction
temp1$pv_S_lam10_lam20_adj_NW #P value of it
```

recursive_hstep_fast *Forecasting h-steps ahead using Recursive Least Squares Fast*

Description

Consider the following LS-fitted Model with intercept: $y_{(t+h)} = \beta_0 + x_t * \beta + u_{(t+h)}$ which is used to generate out-of-sample forecasts of y , h -steps ahead ($h=1,2,3,\dots$). It calculates the recursive residuals starting from the first ($n * \pi_0$) data points, where n is the total number of data points.

Usage

```
recursive_hstep_fast(y, x, pi0, h)
```

Arguments

<code>y</code>	$n \times 1$ Outcome series, which should be numeric and one dimensional.
<code>x</code>	$n \times p$ Predictor matrix (intercept would be added automatically).
<code>pi0</code>	Fraction of the sample, which should be within 0 and 1.
<code>h</code>	Number of steps ahead to predict, which should be a positive integer.

Details

`recursive_hstep_fast` is the fast version that avoids the recursive calculation of inverse of the matrix using Sherman-Morrison formula. `recursive_hstep_slow` is the slow version that calculates the standard OLS recursively.

Value

Series of residuals estimated

Author(s)

Rong Peng, <r.peng@soton.ac.uk>

Examples

```
x<- rnorm(15);
y<- x+rnorm(15);
temp1 <- recursive_hstep_fast(y,x,pi0=0.5,h=1);
```

recursive_hstep_slow *Forecasting h-steps ahead using Recursive Least Squares Slow*

Description

Consider the following LS-fitted Model with intercept: $y_{(t+h)} = \beta_0 + x_t * \beta + u_{(t+h)}$ which is used to generate out-of-sample forecasts of y , h -steps ahead ($h=1,2,3,\dots$). It calculates the recursive residuals starting from the first ($n * \pi_0$) data points, where n is the total number of data points.

Usage

```
recursive_hstep_slow(y, x, pi0, h)
```

Arguments

<code>y</code>	$n \times 1$ Outcome series, which should be numeric and one dimensional.
<code>x</code>	$n \times p$ Predictor matrix (intercept would be added automatically).
<code>pi0</code>	Fraction of the sample, which should be within 0 and 1.
<code>h</code>	Number of steps ahead to predict, which should be a positive integer.

Details

`recursive_hstep_fast` is the fast version that avoids the recursive calculation of inverse of the matrix using Sherman-Morrison formula. `recursive_hstep_slow` is the slow version that calculates the standard OLS recursively.

Value

Series of residuals estimated

Author(s)

Rong Peng, <r.peng@soton.ac.uk>

Examples

```
x<- rnorm(15);
y<- x+rnorm(15);
temp2 <- recursive_hstep_slow(y,x,pi0=0.5,h=1);
```

Index

dm_cw, [2](#)

lr_var, [2](#)

Nested_Stats_S0, [3](#), [5](#)

Nested_Stats_Sbar, [4](#), [4](#)

recursive_hstep_fast, [6](#)

recursive_hstep_slow, [7](#)