

Package ‘RATest’

October 11, 2017

Type Package

Title Randomization Tests

Description A collection of randomization tests, data sets and examples. The current version focuses on the description and implementation of a permutation test for testing the continuity assumption of the baseline covariates in the sharp regression discontinuity design (RDD) as in Canay and Kamat (2017) <<https://goo.gl/UZFqt7>>. More specifically, it allows the user to select a set of covariates and test the aforementioned hypothesis using a permutation test based on the Cramer-von Miss test statistic. Graphical inspection of the empirical CDF and histograms for the variables of interest is also supported in the package.

Version 0.1.1

Maintainer Mauricio Olivares-Gonzalez <maurice.olivares@gmail.com>

Depends R (>= 3.00), ggplot2(>= 2.2.1), gridExtra

Imports stats, quantreg

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Author Mauricio Olivares-Gonzalez [aut, cre],
Ignacio Sarmiento-Barbieri [aut]

Collate 'CvM.stat.R' 'H.cdf.R' 'RDperm.R' 'lee2008.R' 'plot.RDperm.R'
'summary.RDperm.R'

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2017-10-11 18:15:25 UTC

R topics documented:

| | |
|-----------------------|---|
| CvM.stat | 2 |
| H.cdf | 3 |
| lee2008 | 4 |
| plot.RDperm | 4 |
| RDperm | 5 |

| | |
|--------------|----------|
| Index | 8 |
|--------------|----------|

| | |
|----------|--------------------------------------|
| CvM.stat | <i>Cramer - von Mises statistics</i> |
|----------|--------------------------------------|

Description

Calculates the Cramer-von Mises test statistic

$$T(S_n) = \frac{1}{2q} \sum_{i=1}^{2q} (H_n^-(S_{n,i}) - H_n^+(S_{n,i}))^2$$

where $H_n^-(\cdot)$ and $H_n^+(\cdot)$ are the empirical CDFs of the the sample of baseline covariates close to the cutoff from the left and right, respectively. See equation (12) in Canay and Kamat (2017).

Usage

`CvM.stat(Sn)`

Arguments

`Sn` Numeric. The pooled sample of induced order statistics. The first column of `S` can be viewed as an independent sample of `W` conditional on `Z` being close to zero from the left. Similarly, the second column of `S` can be viewed as an independent sample of `W` conditional on `Z` being close to the cutoff from the right. See section 3 in Canay and Kamat (2017).

Value

Returns the numeric value of the Cramer - von Mises test statistic.

Author(s)

Mauricio Olivares Gonzalez
Ignacio Sarmiento Barbieri

References

Canay, I and Kamat V, (2017) Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design. <http://faculty.wcas.northwestern.edu/~iac879/wp/RDDPermutations.pdf>

Description

Calculates the empirical CDF of the sample of W conditional on Z being close to the cutoff from either the left or right. Given the induced order for the baseline covariates

$$W_{[q]}^-, W_{[q-1]}^-, \dots, W_{[1]}^-$$

or

$$W_{[1]}^+, W_{[2]}^+, \dots, W_{[q]}^+$$

, this function will calculate either

$$H_n^-(t) = \frac{1}{q} \sum_{i=1}^q I\{W_{[i]}^- \leq t\}$$

or

$$H_n^+(t) = \frac{1}{q} \sum_{i=1}^q I\{W_{[i]}^+ \leq t\}$$

depending on the argument of the function. See section 3 in Canay & Kamat (2017).

Usage

H.cdf(W, t)

Arguments

W Numeric. The sample of induced order statistics. The input can be either $\{W_{[q]}^-, W_{[q-1]}^-, \dots, W_{[1]}^-\}$ or $\{W_{[1]}^+, W_{[2]}^+, \dots, W_{[q]}^+\}$.

t Numeric. The scalar needed for the calculation of the CDF.

Value

Numeric. For a sample $W = (w_1, \dots, w_n)$, returns the fraction of observations less or equal to t .

Author(s)

Mauricio Olivares Gonzalez

Ignacio Sarmiento Barbieri

References

Canay, I and Kamat V, (2017) Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design. <http://faculty.wcas.northwestern.edu/~iac879/wp/RDDPermutations.pdf>

lee2008

Dataset used in Lee (2008)

Description

Randomized experiments from non-random selection in U.S. House elections

Format

A data frame with 6558 observations and two variables:

demsharenext Democrat vote share election t+1

difdemshare Running variable. Diff. democratic share

demshareprev Democrat vote share t-1

demwinprev Democrat win t-1

demofficeexp Democrat political experience t

othofficeexp Oppositions political experience t

demelectexp Democrat electoral experience t

othelectexp Opposition electoral experience t

Source

Mostly Harmless Econometrics Data Archive: <http://economics.mit.edu/faculty/angrist/data1/mhe>

References

Lee, D. (2008) Randomized experiments from non-random selection in U.S. House elections, *Journal of Econometrics*, 142, 675-697

plot.RDperm*Plot RDperm*

Description

Plots a histogram and empirical cdf

Usage

```
## S3 method for class 'RDperm'  
plot(x, w, plot.class = "both", ...)
```

Arguments

| | |
|------------|---|
| x | Object of class "RDperm" |
| w | Character. Name of variable to be plotted |
| plot.class | Character. Can be: "both" for a histogram and cdf plot, "hist" for a histogram or "cdf" for only the cdf plot |
| ... | Additional ggplot2 controls |

Author(s)

Mauricio Olivares Gonzalez

Ignacio Sarmiento Barbieri

References

Canay, I and Kamat V, (2017) Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design. <http://faculty.wcas.northwestern.edu/~iac879/wp/RDDPermutations.pdf>

Examples

```
## Not run:
permtest<-RDperm(W=c("demshareprev","demwinprev"),z="difdemshare",data=lee2008)
plot(permtest,w="demshareprev")

## End(Not run)
```

RDperm

Regression Discontinuity Design Permutation test

Description

A permutation test for continuity of covariates in Sharp Regression Discontinuity Design as described in Canay and Kamat (2017).

Usage

```
RDperm(W, z, data, n.perm = 499, q_type = 10, cutoff = 0,
  test.statistic = "CvM")
```

Arguments

| | |
|---|---|
| W | Character. Vector of covariates names. The procedure will test the null hypothesis of continuity of the distribution of each element in W at the cutoff. |
| z | Character. Running variable name. This is the scalar random variable that defines, along with the cutoff, the treatment assignment rule in the sharp regression discontinuity design. |

| | |
|-----------------------------|---|
| <code>data</code> | Data.frame. |
| <code>n.perm</code> | Numeric. Number of permutations needed for the stochastic approximation of the p-values. See remark 3.2 in Canay and Kamat (2017). The default is B=499. |
| <code>q_type</code> | A fixed and small (relative to the sample size) natural number that will define the q closest values of the order statistic of Z to the right and to the left of the cutoff. If 'rot', it calls for the Rule of Thumb described in Canay and Kamat (2017), section 3.1. |
| <code>cutoff</code> | Numeric. The scalar defining the threshold of the running variable. |
| <code>test.statistic</code> | Character. A rank test statistic satisfying rank invariance. The default is a Cramer-von Mises test statistic. |

Value

The functions `summary` and `plot` are used to obtain and print a summary and plot of the estimated regression discontinuity. The object of class `RDperm` is a list containing the following components:

| | |
|-----------------------------|--|
| <code>results</code> | Matrix. Test Statistic, P-values and Q |
| <code>test.statistic</code> | Test Statistic |
| <code>q_type</code> | Type of Q used in the calculations, can be either, "Defined by User", the "Rule of Thumb" ("rot") or the "Alternative Rule of Thumb" ("arot). If the user wants to specify the q, he should supply a number, the default is 10. In the case the user wants to use the "Rule of Thumb" or the "Alternative Rule of Thumb", he/she should specify "rot" or "arot". |
| <code>n_perm</code> | number of permutations |
| <code>rv</code> | Character. Running variable name |
| <code>Z</code> | Vector. Running Variable |
| <code>cutoff</code> | cutoff |
| <code>data</code> | data set |
| <code>S</code> | Matrix. Pooled sample of induced order statistics |
| <code>S_perm</code> | List. Permutations of the induced order statistic. |

Author(s)

Mauricio Olivares Gonzalez

Ignacio Sarmiento Barbieri

References

Canay, I and Kamat V, (2017) Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design. <http://faculty.wcas.northwestern.edu/~iac879/wp/RDDPermutations.pdf>

Examples

```
permtest<-RDperm(W=c("demshareprev"),z="difdemshare",data=lee2008)
summary(permtest)
## Not run:
permtest<-RDperm(W=c("demshareprev","demwinprev"),z="difdemshare",data=lee2008)
summary(permtest)

## End(Not run)
```

Index

*Topic **permutation**

CvM.stat, 2

H.cdf, 3

plot.RDperm, 4

RDperm, 5

*Topic **rdperm**

CvM.stat, 2

H.cdf, 3

plot.RDperm, 4

RDperm, 5

*Topic **test**

CvM.stat, 2

H.cdf, 3

RDperm, 5

CvM.stat, 2

H.cdf, 3

lee2008, 4

plot.RDperm, 4

RDperm, 5