Package 'ThesisSourceCode'

April 7, 2023

Type Package

Title Regression Discontinuity Design Simulation of Different Covariate Selection Procedures

Version 1.0

RoxygenNote 7.2.3

R topics documented:

`calculate_correlation` *Calculate the sample correlation coefficient of two vectors*

Description

Calculate the sample correlation coefficient of two vectors

Usage

```
calculate_correlation(A, B)
```

Arguments

A	Vector of dimension n
B	Vector of dimension n

Value

Sample correlation coefficient of A and B

`calculate_correlation_thresholds`
Calculate correlation thresholds for each covariate with the outcome according to Section 8.3.1 of the thesis

---

Description

Calculate correlation thresholds for each covariate with the outcome according to Section 8.3.1 of the thesis

Usage

```
calculate_correlation_thresholds(Z, Y, data_size)
```

Arguments

`Z`	The covariate matrix of dimension nxp
`Y`	The outcome vector of dimension n
`data_size`	The sample size n

Value

The threshold vector containing a threshold for each of the covariates

---

`calculate_correlation_threshold_matrix`
Calculate deletion correlation thresholds for each pair of covariates according to Section 8.3.2 of the thesis

---

Description

Calculate deletion correlation thresholds for each pair of covariates according to Section 8.3.2 of the thesis

Usage

```
calculate_correlation_threshold_matrix(Z, data_size)
```

Arguments

`Z`	A matrix containing the selected covariates by the selection procedure of Section 8.3.1 of dimension nxs
`data_size`	The sample size n

Value

The matrix containing the thresholds for each pair of covariates

compare_correlation — *Calculate the sample correlation coefficient of each covariate with the outcome and check whether it is greater or equal a threshold*

Description

Calculate the sample correlation coefficient of each covariate with the outcome and check whether it is greater or equal a threshold

Usage

```
compare_correlation(Z, Y, threshold)
```

Arguments

Z	The covariates matrix of dimension nxp
Y	The outcome vector of dimension n
threshold	The threshold vector of dimension n

Value

Returns indices of covariates which have a sample correlation to the outcome greater or equal the according threshold

cross_interactions — *Compute cross-interaction terms of two sets of covariates*

Description

Compute cross-interaction terms of two sets of covariates

Usage

```
cross_interactions(Z1, Z2, ident = "")
```

Arguments

Z1	The covariate matrix 1 of dimension n x p1
Z2	The covariate matrix 2 of dimension n x p2
ident	String for naming the resulting columns

Value

All cross-interaction terms structured as a matrix

fourier_basis *Calculate trigonometric transformations of the covariates*

Description

Calculate trigonometric transformations of the covariates

Usage

```
fourier_basis(Z, order)
```

Arguments

Z	The covariate matrix of dimension nxp
order	The order up to which the transformations should be calculated

Value

The trigonometric transformations structured as a matrix

interaction_terms *Compute interaction terms of covariates*

Description

Compute interaction terms of covariates

Usage

```
interaction_terms(Z)
```

Arguments

Z	The covariate matrix of dimension nxp

Value

All interaction terms structured as matrix

`perform_rdd` *Execute a RD analysis on generated data including covariates chosen by different selection procedures (used to generate results of Section 9.2.1 and 9.2.2)*

Description

Execute a RD analysis on generated data including covariates chosen by different selection procedures (used to generate results of Section 9.2.1 and 9.2.2)

Usage

```
perform_rdd(run, sample_size, rdd_library = "honest", estimator_type = 1)
```

Arguments

`run`	The number of execution when executed multiple times in parallel
`sample_size`	The sample size
`rdd_library`	The R package to use for the RD analysis. Possible values are "honest" for the package RDHonest and "robust" for the package RDRobust.
`estimator_type`	This parameter is just relevant when using RDRobust (otherwise it can be ignored). It indicates the estimator type used in the RD analysis. Possible values are: 1 - for conventional estimator, 2 - for bias-corrected estimator, 3 - for robust estimator

Value

The results of the RD analysis (estimation, bias, standard deviation, standard error, confidence intervals, coverage) and the results on the selection of covariates

`perform_rdd_on_data` *Execute a RD analysis on a given data set including covariates chosen by different selection procedures (used to generate results of Section 9.4)*

Description

Execute a RD analysis on a given data set including covariates chosen by different selection procedures (used to generate results of Section 9.4)

Usage

```
perform_rdd_on_data(X, Y, Z, rdd_library = "honest", estimator_type = "1")
```

Arguments

X	The running variable array of dimension n
Y	The outcome array of dimension n
Z	The covariate matrix of dimension nxp
rdd_library	The R package to use for the RD analysis. Possible values are "honest" for the package RDHonest and "robust" for the package RDRobust.
estimator_type	This parameter is just relevant when using RDRobust (otherwise it can be ignored). It indicates the estimator type used in the RD analysis. Possible values are: 1 - for conventional estimator, 2 - for bias-corrected estimator, 3 - for robust estimator

Value

The results of the RD analysis (estimation, standard error, confidence intervals) as well as the results on the selection of covariates

---

perform_rdd_redundant_covariates

Execute a RD analysis on generated data including covariates chosen from a set of redundant covariates by different selection procedures (used to generate results of Section 9.3)

Description

Execute a RD analysis on generated data including covariates chosen from a set of redundant covariates by different selection procedures (used to generate results of Section 9.3)

Usage

```
perform_rdd_redundant_covariates(
  run,
  sample_size,
  rdd_library = "honest",
  estimator_type = 1
)
```

Arguments

run	The number of execution when executed multiple times in parallel
sample_size	The sample size
rdd_library	The R package to use for the RD analysis. Possible values are "honest" for the package RDHonest and "robust" for the package RDRobust.
estimator_type	This parameter is just relevant when using RDRobust (otherwise it can be ignored). It indicates the estimator type used in the RD analysis. Possible values are: 1 - for conventional estimator, 2 - for bias-corrected estimator, 3 - for robust estimator

Value

The results of the RD analysis (estimation, bias, standard deviation, standard error, confidence intervals, coverage) and the results on the selection of covariates

`remove_covs_calculated_threshold` *Remove selected covariates according to the procedure described in Section 8.3.2 and 8.3.3 (simple and advanced deletion, respectively)*

Description

Remove selected covariates according to the procedure described in Section 8.3.2 and 8.3.3 (simple and advanced deletion, respectively)

Usage

```
remove_covs_calculated_threshold(Z, Y, data_size, simple_deletion = TRUE)
```

Arguments

`Z`	A matrix containing the selected covariates by the Selection procedure of Section 8.3.1 of dimension nxs
`Y`	The outcome vector of dimension of n
`data_size`	The sample size n
`simple_deletion`	When set to TRUE, simple deletion is applied (Section 8.3.2), otherwise advanced deletion (Section 8.3.3)

Value

The remaining covariates after deletion

`remove_covs_with_high_correlation` *Removes a certain amount of covariates with the highest sample correlation to another covariate. This function is just used to ensure invertibility in some cases as described in the paragraph "Numerical invertibility" of Section 9.1*

Description

Removes a certain amount of covariates with the highest sample correlation to another covariate. This function is just used to ensure invertibility in some cases as described in the paragraph "Numerical invertibility" of Section 9.1

Usage

```
remove_covs_with_high_correlation(Z, number)
```

Arguments

`Z`	The covariate matrix of dimension nxp
`number`	The number of covariates to be deleted

Value

The remaining covariates after deletion

`triangular` *Compute the triangular kernel*

Description

Compute the triangular kernel

Usage

```
triangular(x)
```

Arguments

`x`	The argument passed to the triangular kernel function.

Value

The value of the triangular kernel function.

Index