Package 'fastrerandomize'

October 2, 2023

Title fastrerandomize: R Package for Ultra-fast Re-randomization Using a JAX Backend

Version 0.1

Author 'Connor Jerzak <connor.jerzak@gmail.com> [aut, cre], Rebecca Goldstein <rgoldstein@berkeley.edu> [aut]'

Description An R Package for Ultra-fast Re-randomization Using a JAX Backend

Depends R (>= 3.3.3)

License Creative Commons Attribution-Noncommercial-No Derivative Works 4.0, for academic use only.

Encoding UTF-8

LazyData false

Maintainer 'Connor Jerzak' <connor.jerzak@gmail.com>

Imports reticulate

RoxygenNote 7.2.3

R topics documented:

`GenerateCausalData` *This function generates simulated causal data based on specified parameters.*

Description

This function generates simulated causal data based on specified parameters.

Usage

```
GenerateCausalData(n_units)
```

Arguments

`n_units`	A numeric value specifying the total number of units in the sample.
`proportion_treated`	A numeric value between 0 and 1 indicating the proportion of units that receive treatment.
`k_covars`	A numeric value indicating the number of covariates to be generated.
`rho`	A numeric value representing the correlation coefficient.
`SD_inherent`	A numeric value indicating the standard deviation inherent to the data.
`treatment_effect_mean`	A numeric value representing the mean of the treatment effect.
`treatment_effect_SD`	A numeric value indicating the standard deviation of the treatment effect.
`covariates_SD`	A numeric value or vector specifying the standard deviation of the covariates.
`Y0_coefficients`	An optional numeric vector specifying the coefficients for the control outcome model. If not provided, the function assumes a NULL value.
`Y1_coefficients`	An optional numeric vector specifying the coefficients for the treated outcome model. If not provided, the function assumes a NULL value.

Value

A list consisting of

- `data_matrix` A data frame containing the simulated covariates and outcomes for both control (Y0) and treatment (Y1) groups.
- `Y0_coefficients` A numeric vector representing the coefficients used for the control outcome model.
- `Y1_coefficients` A numeric vector representing the coefficients used for the treated outcome model.

Examples

```
# For a tutorial, see
# github.com/cjerzak/fastrerandomization-software
```

`GenerateRandomizations` *Fast generation of all possible complete randomizations given target number of experimental units.*

Description

Fast generation of all possible complete randomizations given target number of experimental units.

Usage

```
GenerateRandomizations(n_units, n_treated)
```

Arguments

n_units	A integer specifying total number of experimental units.
n_treated	An integer specifying total number of treated units.

Value

A JAX array containing all possible complete randomizations.

Examples

```
# For a tutorial, see
# github.com/cjerzak/fastrerandomization-software
```

InitializeJAX *Initialize JAX*

Description

Initialize JAX

Usage

```
InitializeJax(conda_env, conda_env_required)
```

Arguments

conda_env	An optioanl character string representing the conda environment to activate. A version of JAX should live in that environment. If NULL, we look in the default Python environment for JAX.
conda_env_required	A logical representing whether to force use the specified conda environment. Used only if conda_env specified.

Value

This function initializes a JAX-containing conda environment as specified by conda_env. This function must be run before any others in fastrerandomize.

Examples

```
# For a tutorial, see
# github.com/cjerzak/fastrerandomization-software
```

QJEData *QJEData*

Description

The dataset originates from the study "Moral hazard: Experimental evidence from tenancy contracts" by Burchardi, Konrad B et al., published in "The Quarterly Journal of Economics" in 2019 (Volume 134, Issue 1, Pages 281-347).

Usage

```
QJEData
```

Format

A data frame with 968 rows and many columns containing treatment data for a Quarterly Journal of Economics experiment on agriculture.

Source

Burchardi, Konrad B et al. (2019). "Moral hazard: Experimental evidence from tenancy contracts." In: The Quarterly Journal of Economics 134.1, pp. 281–347

RandomizationTest *Fast randomization test*

Description

Fast randomization test

Usage

```
RandomizationTest(X, ...)
```

Arguments

obsW	A numeric vector where 0's correspond to control units and 1's to treated units.
obsY	An optional numeric vector of observed outcomes. If not provided, the function assumes a NULL value.
X	A numeric matrix of covariates.
alpha	The significance level for the test. Default is 0.05.
candidate_randomizations	A numeric matrix of candidate randomizations.
candidate_randomizations_array	An optional JAX array of candidate randomizations. If not provided, the function coerces candidate_randomizations into a JAX array.
n0_array	An optional array specifying the number of control units.
n1_array	An optional array specifying the number of treated units.

`prior_treatment_effect_mean`
: An optional numeric value for the prior mean of the treatment effect. Default is NULL.

`prior_treatment_effect_SD`
: An optional numeric value for the prior standard deviation of the treatment effect. Default is NULL.

`true_treatment_effect`
: An optional numeric value specifying the true treatment effect. Default is NULL.

`simulate`
: A logical value indicating whether to run `RandomizationTest` in simulation mode. Default is FALSE.

`coef_prior`
: An optional function generating coefficients on values of `X` for predicting `Y(0)`.

`nSimulate_obsW`
: A numeric value specifying the number of simulated values for obsW. Default is `50L`.

`nSimulate_obsY`
: A numeric value specifying the number of simulated values for obsY. Default is `50L`.

`randomization_accept_prob`
: An numeric scalar or vector of probabilities for accepting each randomization.

`findFI`
: A logical value indicating whether to find the fiducial interval. Default is FALSE.

`c_initial`
: A numeric value representing the initial criterion for the randomization. Default is `2`.

Value

A list consisting of

- `p_value` A numeric value or vector representing the p-value of the test (or the expected p-value under the prior structure specified in the function inputs).
- `FI` A numeric vector representing the fiducial interval if `findFI=T`.
- `tau_obs` A numeric value or vector representing the estimated treatment effect(s)

References

-

Examples

```
# For a tutorial, see
# github.com/cjerzak/fastrerandomization-software
```

Index