# Package: twangRDC (via r-universe)

September 13, 2024

**Version** 1.0

**Date** 2021-05-10

**Title** Gradient Boosting for Linkage Failure in FSRDCs

**Maintainer** Matthew Cefalu <Matthew_Cefalu@rand.org>

**Depends** R (>= 3.5.0), xgboost, data.table, ggplot2

**Imports** MatrixModels

**Suggests** knitr, rmarkdown

**Description** Provides functions for gradient boosted weighting to correct linkage failures or generate comparison groups.

**License** GPL-3

**Encoding** UTF-8

**VignetteBuilder** knitr, rmarkdown

**RoxygenNote** 7.1.1

**NeedsCompilation** no

**Author** Matthew Cefalu [aut, cre], John Sullivan [aut]

**Date/Publication** 2021-05-17 21:40:02 UTC

**Repository** https://mattcefalu.r-universe.dev

**RemoteUrl** https://github.com/cran/twangRDC

**RemoteRef** HEAD

**RemoteSha** 3065f4e5d54c585abe90c427c89ed5975c4e5365

## Contents

bal.table *Balance tables for* twangRDC

## Description

bal.table calculates balance tables from a ps.xgb object

## Usage

```
bal.table(
  x,
  type = "overall",
  n = 10,
  decreasing = TRUE,
  which.sort = "adj",
  include.var = FALSE
)
```

## Arguments

| | |
|---|---|
| x | A ps.xgb object. |
| type | An optional character string requesting if balance should be summarized overall ("overall") or by strata ("strata"). Default: "overall". |
| n | An integer specifying the number of rows to print in the balance table. Default: 10. |
| decreasing | A logical value indicating if the balance table should be sorted in increasing or decreasing order. |
| which.sort | An optional character string indicating if the balance table should be sorted by the adjusted ("adj") or unadjusted ("unadj") absolute standardized differences. Default: "adj". |
| include.var | A logical value indicating if the variable corresponding the the maximum absolute standardized difference within strata should be included in the balance table. Only valid when strata=TRUE. Default: FALSE. |

## Value

Returns a table.

## Examples

```
# See vignette for examples.
```

## get.weights *Exact weights from a ps.xgb object*

### Description

Extracts weights from a ps.xgb object, output with the unique identifier for easy merging.

### Usage

```
get.weights(x)
```

### Arguments

| | |
|---|---|
| x | An object of class ps.xgb |

### Value

Returns a data frame.

### Examples

```
# See vignette for examples.
```

## nola_south *An example FSRDC dataset*

### Description

See vignette for detailed description of the data.

### Usage

```
data(nola_south)
```

### Format

A data frame

- tract_id_str. Census tract identifier
- metarea. metropolitan area
- c00_age12. categorical variable for age
- c00_sex. sex
- c00_race. categorical variable for race
- c00_nphu. number of persons in housing unit

- hhid. household identifier
- sim_pik. indicator of PIK assignment
- nola_rec. indicator of record from Orleans Parish
- id. individual identifier

A data frame with 18396 rows and 10 variables

---

`plot.ps.xgb` *Plot for a ps.xbg object*

---

## Description

`ps.xgb.plot` produces a figure showing the balance criteria by iteration for a ps.xgb object.

## Usage

```
## S3 method for class 'ps.xgb'
plot(x, ...)
```

## Arguments

| | |
|---|---|
| `x` | An object of class ps.xgb |
| `...` | Arguments to be passed to other functions |

## Value

Returns a ggplot object.

## Examples

```
# See vignette for examples.
```

---

`ps.xgb` *Gradient boosted propensity score estimation*

---

## Description

`ps.xgb` calculates propensity scores using gradient boosted logistic regression and diagnoses the resulting propensity scores using a variety of methods

## Usage

```
ps.xgb(
  formula = formula(data),
  strata = NULL,
  data,
  params,
  file = NULL,
  max.steps = Inf,
  iters.per.step = 100,
  id.var,
  min.iter = 1000,
  min.width = NULL,
  verbose = TRUE,
  save.model = FALSE,
  weights = NULL,
  linkage = TRUE
)
```

## Arguments

| | |
|---|---|
| formula | An object of class formula: a symbolic description of the propensity score model to be fit with the treatment indicator on the left side of the formula and the variables to be balanced on the right side. |
| strata | An optional factor variable identifying the strata. If specified, balance is optimized within strata. |
| data | A dataset. |
| params | xgboost parameters. |
| file | An optional character string naming a file to save intermediate results. |
| max.steps | An integer specifying the maximum number of steps to take. Note that max.steps*iters.per.step must be greater than or equal to min.iter. Default: Inf. |
| iters.per.step | An integer specifying the number of iterations to add to the model at each step of algorithm. Note that max.steps*iters.per.step must be greater than or equal to min.iter. Default: 100. |
| id.var | A variable that uniquely identifies observations. |
| min.iter | An integer specifying the minimum number of iterations before checking for convergence. Note that max.steps*iters.per.step must be greater than or equal to min.iter. Default: 1000. |
| min.width | An integer specifying the minimum number of iterations between the current number of iterations and the optimal value. Default: 5*iters.per.step. |
| verbose | A logical value indicating if the function should update the user on its progres Default: TRUE. |
| save.model | A logical value indicating if the xgboost model be saved as part of the output object. Default: FALSE. |
| weights | An optional variable that identifies user defined weights to be incorporated into the optimization. |

`linkage` An indicator of whether the weighting should be for linkage failure (or non-response) versus comparison group construction. A value of TRUE requests weighting to account for linkage failure, while a value of FALSE requests weighting for comparison group construction. Default: TRUE.

## Value

Returns an object of class `ps.xgb`, a list containing

- `bal.tab` A table summarizing the balance at the optimal number of iterations.
- `es` A table summarizing the standardized differences within strata at the optimal number of iterations.
- `es.max` A table summarizing the maximum absolute standardized difference by strata.
- `es.mean`A table summarizing the mean absolute standardized difference by strata.
- `iter.per.step` Saves the value of `iters.per.step` specified by the user.
- `opt.iter` The optimal number of iterations.
- `strata` A list of the strata used in the optimization.
- `weight.data` A dataset containing the unique ID and the optimal weight for each observation.

## References

Dan McCaffrey, G. Ridgeway, Andrew Morral (2004). "Propensity Score Estimation with Boosted Regression for Evaluating Adolescent Substance Abuse Treatment", *Psychological Methods* 9(4):403-425.

## See Also

twang::ps, xgboost

## Examples

```
# See the vignette for examples.
```

# Index