# Package 'twangMediation'

July 26, 2021
Title Twang Causal Mediation Modeling via Weighting
Version 1.0
Author Dan McCaffrey [aut, cre], Katherine Castellano [aut], Donna Coffman [aut], Brian Vegetabile [aut], Megan Schuler [aut], Haoyu Zhou [aut]
Maintainer Dan McCaffrey <pre><dmccaffrey@ets.org></dmccaffrey@ets.org></pre>
Description Provides functions for estimating natural direct and indirect effects for mediation analysis. It uses weighting where the weights are functions of estimates of the probability of exposure or treatment assignment (Hong, G (2010). <a href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf">https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf</a> Huber, M. (2014). <a href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf">https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf</a> Huber, M. (2014). <a a="" cepa.stanford.edu="" default="" files="" gh_jsm%20proceedings%202010.pdf<="" href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf&lt;/a&gt; Huber, M. (2014). &lt;a href=" https:="" sites="" workshops=""> Huber, M. (2014). <a a="" cepa.stanford.edu="" default="" files="" gh_jsm%20proceedings%202010.pdf<="" href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf&lt;/a&gt; Huber, M. (2014). &lt;a href=" https:="" sites="" workshops=""> Huber, M. (2014). <a a="" cepa.stanford.edu="" default="" files="" gh_jsm%20proceedings%202010.pdf<="" href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf&lt;/a&gt; Huber, M. (2014). &lt;a href=" https:="" sites="" workshops=""> Huber, M. (2014). <a a="" cepa.stanford.edu="" default="" files="" gh_jsm%20proceedings%202010.pdf<="" href="https://cepa.stanford.edu/sites/default/files/workshops/GH_JSM%20Proceedings%202010.pdf&lt;/a&gt; Huber, M. (2014). &lt;a href=" https:="" sites="" workshops=""> Huber Huber, M. (2014). </a></a></a></a>

2 bal.table.mediation

desc.effects.mediation
dx.wts.mediation
NSDUH_female
plot.mediation
print.bal.table.mediation
print.mediation
summary.mediation
swapTxCtrl
tMdat
weighted_mean
wgtmed

twangMediation-package

twangMediation: Twang Causal Mediation Modeling via Weighting

**19** 

## **Description**

Index

Provides functions for estimating natural direct and indirect effects for mediation analysis. It uses weighting where the weights are functions of estimates of the probability of exposure or treatment assignment (Hong, G (2010). <a href="https://cepa.stanford.edu/sites/default/files/workshops/GH\_JSM">https://cepa.stanford.edu/sites/default/files/workshops/GH\_JSM</a>

## Author(s)

Maintainer: Dan McCaffrey <dmccaffrey@ets.org>

Authors:

- Katherine Castellano < kecastellano@ets.org>
- Donna Coffman <donna.coffman@temple.edu>
- Brian Vegetabile <br/>bvegetab@rand.org>
- Megan Schuler <mschuler@rand.org>
- Haoyu Zhou <haoyu.zhou@temple.edu>

bal.table.mediation Compute the balance table for mediation object.

## Description

Compute the balance table for mediation object.

## Usage

```
bal.table.mediation(x, digits = 3, ...)
```

calculate\_effects 3

## **Arguments**

x A mediation object

digits Number of digits to round to. Dafault: 3

... Additional arguments.

#### Value

res tables detailing covariate balance across exposure groups both before and after

weighting

## See Also

```
print.bal.table.mediation, wgtmed
```

## **Examples**

calculate\_effects

Calculate the actual effects

#### **Description**

Calculate the actual effects

## Usage

```
calculate_effects(w_11, w_00, w_10, w_01, y_outcome, sampw = NULL)
```

## **Arguments**

w_11	The $Y(1, M(1))$ weights
w_00	The $Y(0, M(0))$ weights
w_10	The $Y(1, M(0))$ weights
w_01	The $Y(0, M(1))$ weights

y\_outcome The Y variable

sampw Sampling weights, set to NULL by default.

4 desc.effects

#### Value

res

The actual effects

#### See Also

wgtmed

check\_missing

Check vector for NA or NAN values.

## Description

check\_missing raises and error if the data contains. NA or NAN values.

## Usage

```
check_missing(x)
```

#### **Arguments**

Х

numeric The data set to check for NA or NAN values.

#### Value

Indicator of the existence of NA or NAN values

desc.effects

Describe the effects

## Description

Describe the effects, and calculate standard errors and confidence intervals

## Usage

```
desc.effects(x, ...)
```

## **Arguments**

x An object

... list, optional Additional arguments.

## Value

Effects, standard errors and confidence intervals of an object

#### See Also

```
{\tt desc.effects.mediation}, {\tt wgtmed}
```

desc.effects.mediation 5

#### **Examples**

desc.effects.mediation

Describe the effects from a mediation object

## Description

Describe the effects, and calculate standard errors and confidence intervals from a mediation object

## Usage

```
## S3 method for class 'mediation'
desc.effects(x, y_outcome = NULL, ...)
```

#### **Arguments**

x A mediation object

y\_outcome The outcome; if NULL, then Y must have been provided to the original mediation

function.

... Additional arguments..

#### Value

results effects, standard errors, and confidence intervals of a mediation object

## See Also

```
desc.effects, wgtmed
```

6 dx.wts.mediation

#### **Examples**

dx.wts.mediation

Compute diagnostics assessing covariates balance.

## **Description**

dx.wts.mediation takes a ps object or a set of propensity scores and computes diagnostics assessing covariates balance.

## Usage

```
dx.wts.mediation(
    x,
    data,
    estimand,
    vars = NULL,
    treat.var,
    x.as.weights = TRUE,
    sampw = NULL,
    perm.test.iters = 0
```

## **Arguments**

X	A data frame, matrix, or vector of propensity score weights or a ps object. x can
	also be a data frame, matrix, or vector of propensity scores if x.as.weights=FALSE.

data A data frame.

estimand The estimand of interest: either "ATT" or "ATE".

vars A vector of character strings naming variables in data on which to assess bal-

ance.

treat.var A character string indicating which variable in data contains the 0/1 treatment

group indicator.

dx.wts.mediation 7

x.as.weights

TRUE or FALSE indicating whether x specifies propensity score weights or propensity scores. Ignored if x is a ps object. Default: TRUE.

sampw

Optional sampling weights. If x is a ps object, then the sampling weights should have been passed to ps and not specified here. dx.wts.mediation will issue a warning if x is a ps object and sampw is also specified.

perm.test.iters

A non-negative integer giving the number of iterations of the permutation test for the KS statistic. If perm.test.iters=0, then the function returns an analytic approximation to the p-value. This argument is ignored is x is a ps object. Setting perm.test.iters=200 will yield precision to within 3% if the true p-value is 0.05. Use perm.test.iters=500 to be within 2%.

#### **Details**

Creates a balance table that compares unweighted and weighted means and standard deviations, computes effect sizes, and KS statistics to assess the ability of the propensity scores to balance the treatment and control groups.

#### Value

Returns a list containing

• treat The vector of 0/1 treatment assignment indicators.

#### See Also

```
wgtmed,bal.table.mediation,print.mediation,summary.mediation
```

#### **Examples**

NSDUH\_female

NSDUH_female	A dataset containing the substance use condition and sexual orientation of 40293 women respondents to the 2017 & 2018 National Survey
	of Drug Use and Health.

## **Description**

A dataset containing the substance use condition and sexual orientation of 40293 women respondents to the 2017 & 2018 National Survey of Drug Use and Health.

## Usage

NSDUH\_female

#### **Format**

A data frame with 40293 rows and 24 variables:

cigmon indiidual smoked any cigarettes within the past month, yes or no

**educ** education level, 1 = less than high school diploma, 2 = high school diploma, 3 = some college/associates degree, 4 = college degree or higher

income level,  $1 \le 20,000, 2 = 20,000 - 449,999, 3 = 50,000 - 70,000, 4 = 75,000 + 70,000, 4 = 100,000 +$ 

NSDUHwt NSDUH sampling weight

vestr NSDUH strata variable

verep NSDUH replicate within stratum

**employ** employment status, 1 = full-time employment, 2 = part-time employment, 3 = student, 4 = unemployed, 5 = other

race 1 = non-Hispanic white, 2 = non-Hispanic Black, 3 = student, 4 = multiracial/other race

alc15 iniciated alcohol use prior to 15 years old

cig15 iniciated smoking prior to 15 years old, yes or no

 $lgb_flag 1 = lesbian, gay or sexual, 0 = heterosexual$ 

alc\_cig\_depend individual meets criteria for either past-year alcohol use disorder or nicotine dependence

weight2y NSDUH sampling weights(scaled for pooling 2017 and 2018 survey years)

```
age age, 1 = 18-25, 2 = 26-34, 3 = 35-49, 4 = 50+
```

#### Value

NSDUH\_female A sample data for demonstration

### **Source**

https://nsduhweb.rti.org/respweb/homepage.cfm

#### See Also

wgtmed

plot.mediation 9

#### **Examples**

```
## Not run:
data(NSDUH_female)
## End(Not run)
```

plot.mediation

Plot the mediation object.

## **Description**

Plot the mediation object.

#### Usage

```
## S3 method for class 'mediation'
plot(
    x,
    plots = "optimize",
    subset = NULL,
    color = TRUE,
    model_subset = NULL,
    ...
)
```

#### **Arguments**

Χ

weighted\_mediation object

plots

An indicator of which type of plot is desired. The options are

- "optimize" A plot of the balance criteria as a function of the GBM iteration.
- "boxplot" Boxplots of the propensity scores for the treatment and control cases
- "es" or "asmd" Plots of the absolute value of the standardized mean difference (effect size) of the pre-treatment variables before and after reweighting
- "density" Distribution plots of NIE1 (distribution of mediator for treatment sample weighted to match distribution of mediator under control for the population) and NIE0 (distribution of mediator for control sample weighted to match distribution of mediator under treatment for the population) for each mediator. For continuous mediators, distributions are plotted with density curves and for categorical (factor) mediators, distributions are plotted with barplots.
- "weights" Histograms of the standardized weights by each stopping rule. Weights are standardized to sum to 1.

subset

Used to restrict which of the stop.methods will be used in the figure. For example subset = c(1,3) would indicate that the first and third stop.methods (in alphabetical order of those specified in the original call to the mediation function) should be included in the figure. If x-method = logistic or crossval, there is no need to subset as there is only one method used.

color If color = FALSE, figures will be gray scale. Default: TRUE.

model\_subset integer Choose either model A (1), model M0 (2), or model M1 (3) only. Argu-

ment is not relevant for plots = density' or weights'.

... Additional arguments.

#### Value

The plot of a mediation object, can be different types.

#### See Also

wgtmed for function input

#### **Examples**

```
print.bal.table.mediation
```

Default print statement for mediation class

## Description

 $Default\ print\ statement\ for\ mediation\ class$ 

## Usage

```
## S3 method for class 'bal.table.mediation' print(x, ...)
```

## **Arguments**

x A bal.table.mediation object.

... Additional arguments.

print.mediation 11

#### Value

Default print statement.

#### See Also

```
bal.table.mediation, wgtmed
```

## **Examples**

print.mediation

Default print statement for mediation class

## **Description**

Default print statement for mediation class

## Usage

```
## S3 method for class 'mediation' print(x, ...)
```

## Arguments

x A mediation object.... Additional arguments.

#### Value

Default print statement.

## See Also

wgtmed for in put.

12 summary.mediation

## **Examples**

summary.mediation

Displays a useful description of a mediation object.

## Description

Displays a useful description of a mediation object.

## Usage

```
## S3 method for class 'mediation'
summary(object, ...)
```

## **Arguments**

```
object A mediation object ... Additional arguments.
```

#### Value

## See Also

wgtmed

swapTxCtrl 13

## **Examples**

swapTxCtrl

Auxiliary function to swap treatment and control in one element of a desc object from a ps object or dx.wts object

## **Description**

Call this in the wgtmed() function and the bal.table.mediation() function.

#### Usage

```
swapTxCtrl(dd)
```

## **Arguments**

dd

numeric An element of a desc object from a ps or dx.wts object

#### Value

A desc object with swapped treatment and control

## See Also

```
ps, dx.wts.mediation
```

14 tMdat

tMdat

Simulated data for twangMediation

## Description

A simulate dataset for demonstrating the functions in the twangMediation package.

## Usage

tMdat

## **Format**

A data frame with 500 rows and 7 variables:

- w1 Simulated continuous covariate
- w2 Simulated continuous covariate
- w3 Simulated continuous covariate
- A Simulated dichotomous exposure indicator
- Y Simulated continuous outcome
- M Simulated mediator that has 11 unique values

te.wgt Estimated inverse probability weight, estimated using GBM via the twang ps function

## Value

tMdat

A sample of simulated data for demonstration

## See Also

wgtmed

## **Examples**

```
## Not run:
data(tMdat)
## End(Not run)
```

weighted\_mean 15

	maan
weighted_	_IIIEan

Calculate a weighted mean.

## **Description**

weighted\_mean calculates a weighted mean, given a vector.

#### Usage

```
weighted_mean(x, weights, multiplier = NULL, na.rm = TRUE)
```

## **Arguments**

```
    x numeric The the data set
    weights numeric The weights
    multiplier An additional vector to multiply Default: NULL
    na.rm Whether to remove NA values. Default: TRUE
```

#### Value

numeric The weighted mean of the data.

wgtmed

Weighted mediation analysis.

## Description

Estimate causal mediation mechanism of a treatment using propensity score weighting.

## Usage

```
wgtmed(
  formula.med,
  data,
  a_treatment,
  y_outcome = NULL,
  total_effect_wts = NULL,
  total_effect_ps = NULL,
  total_effect_stop_rule = NULL,
  method = "ps",
  sampw = NULL,
  ps_n.trees = 10000,
  ps_interaction.depth = 3,
  ps_shrinkage = 0.01,
  ps_bag.fraction = 1,
  ps_n.minobsinnode = 10,
  ps_perm.test.iters = 0,
  ps_verbose = FALSE,
  ps_stop.method = c("ks.mean", "ks.max"),
```

16 wgtmed

```
ps_version = "gbm",
ps_ks.exact = NULL,
ps_n.keep = 1,
ps_n.grid = 25,
ps_cv.folds = 10,
ps_keep.data = FALSE
```

#### **Arguments**

formula.med A object of class formula relating the mediatior(s) to the covariates (potential

confounding variables).

data A dataset of class data.frame that includes the treatment indicator, mediator(s),

and covariates.

a\_treatment The (character) name of the treatment variable, which must be dichotomous (0,

1).

y\_outcome The (character) name of the outcome variable, y. If this is not provided, then no

effects will be calculated and a warning will be raised. Default: NULL.

total\_effect\_wts

A vector of total effect weights, which if left NULL then total\_effect\_ps must be

supplied. Default: NULL.

total\_effect\_ps

A ps object that contains the total effect weights,

total\_effect\_stop\_rule

The stopping rule (ks.mean, ks.max, es.mean, es.max) for the total effect weights, which only needs to be specified if total\_effect\_ps is provided. De-

fault: NULL.

method The method for getting weights ("ps", "logistic", or "crossval"). Default: "ps".

sampw Optional sampling weights Default : NULL.

ps\_n.trees Number of gbm iterations passed on to gbm. Default: 10000.

ps\_interaction.depth

A positive integer denoting the tree depth used in gradient boosting. Default: 3.

ps\_shrinkage A numeric value between 0 and 1 denoting the learning rate. See gbm for more

details. Default: 0.01.

ps\_bag.fraction

A numeric value between 0 and 1 denoting the fraction of the observations randomly selected in each iteration of the gradient boosting algorithm to propose

the next tree. See gbm for more details. Default: 1.0.

 $ps_n.minobsinnode$ 

An integer specifying the minimum number of observations in the terminal nodes of the trees used in the gradient boosting. See gbm for more details.

Default: 10.

ps\_perm.test.iters

A non-negative integer giving the number of iterations of the permutation test for the KS statistic. If perm.test.iters=0 then the function returns an analytic approximation to the p-value. Setting perm.test.iters=200 will yield precision to within 3% if the true p-value is 0.05. Use perm.test.iters=500 to be within 2%. Default: 0.

ps\_verbose

If TRUE, lots of information will be printed to monitor the the progress of the fitting. Default: FALSE.

wgtmed 17

ps\_stop.method A method or methods of measuring and summarizing balance across pretreatment variables. Current options are ks.mean, ks.max, es.mean, and es.max. ks refers to the Kolmogorov-Smirnov statistic and es refers to standardized effect size. These are summarized across the pretreatment variables by either the maximum (.max) or the mean (.mean). Default: c("ks.mean", "ks.max").

ps\_version

"gbm", "xgboost", or "legacy", indicating which version of the twang package to use.

- "gbm" uses gradient boosting from the gbm package.
- "xgboost" uses gradient boosting from the xgboost package.
- "legacy" uses the prior implementation of the ps function.

ps\_ks.exact

NULL or a logical indicating whether the Kolmogorov-Smirnov p-value should be based on an approximation of exact distribution from an unweighted twosample Kolmogorov-Smirnov test. If NULL, the approximation based on the exact distribution is computed if the product of the effective sample sizes is less than 10,000. Otherwise, an approximation based on the asymptotic distribution is used. Warning: setting ks.exact = TRUE will add substantial computation time for larger sample sizes. Default: NULL.

ps\_n.keep

A numeric variable indicating the algorithm should only consider every n. keepth iteration of the propensity score model and optimize balance over this set instead of all iterations. Default: 1.

ps\_n.grid

A numeric variable that sets the grid size for an initial search of the region most likely to minimize the stop.method. A value of n.grid=50 uses a 50 point grid from 1:n. trees. It finds the minimum, say at grid point 35. It then looks for the actual minimum between grid points 34 and 36.If specified with n.keep>1, n.grid corresponds to a grid of points on the kept iterations as defined by n.keep. Default: 25.

ps\_cv.folds

A numeric variable that sets the number of cross-validation folds if using method='crossval'. Default: 10.

ps\_keep.data

A logical variable that determines if the dataset should be saved in the resulting ps model objects. Default: FALSE.

#### **Details**

For users comfortable with ps, any options prefaced with ps\_ are passed directly to the ps() function. Model A is used to estimate  $Pr(A=1 \mid X)$  where X is the vector of background covariates specified in formula.med. If method equals "ps" model A is fit using the twang ps function with estimand= "ATE". If method equals "logistic" then model A is fit using logistic regression. If method equals "crossval" then gbm using cross-validation is used to estimate model A. Because X might include variables not used to estimate the user-provided total effect weights, model A is fit rather than using the user-provided total effect weights to derive  $Pr(A \mid X)$ . If the user uses the same set of variables to estimate their provided total effect weights as they enter in the wgtmed function to estimate the cross-world weights and the user uses the same estimation method and arguments as specified in the wgtmed function, then the estimated model A will match the model the user used to obtain the provided total effect weights.

#### Value

mediation object The mediation object includes the following:

- model\_a The model A ps() results.
- model\_m1 The model M1 ps() results.

18 wgtmed

- model\_m0 The model M0 ps() results.
- data The data set used to compute models
- stopping\_methods The stopping methods passed to stop.method.
- datestamp The date when the analysis was run.
- For each stop.method, a list with the following:
  - TE The total effect.
  - NDE\_0 The natural direct effect, holding the mediator constant at 0.
  - NIE\_1 The natural indirect effect, holding the exposure constant at 1.
  - NDE\_1 The natural direct effect, holding the mediator constant at 1.
  - NIE\_0 The natural indirect effect, holding the exposure constant at 0.
  - expected\_treatment0\_mediator0 E(Y(0, M(0)))
  - expected\_treatment1\_mediator1 E(Y(1, M(1)))
  - expected\_treatment1\_mediator0 E(Y(1, M(0)))
  - expected\_treatment0\_mediator1 E(Y(0, M(1)))

#### See Also

ps

## **Examples**

fit.es.max

```
data("tMdat")
## tMdat is small simulated data set included in twangMediation for
## demonstrating the functions. See ?tMdat for details
head(tMdat)
## The tMdat data contains the following variables:
## w1, w2, w3 -- Simulated covariates
## A -- Simulated dichotomous exposure indicator
## M -- Simulated discrete mediator (11 values)
##
   Y -- Simulated continuous outcome
##
    te.wgt -- Estimated inverse probability weight, estimated using
##
              GBM via the twang ps function
fit.es.max <- wgtmed(M \sim w1 + w2 + w3,
                     data = tMdat,
                      a_treatment = "A",
                      y_outcome = "Y",
                      total_effect_wts = tMdat$te.wgt,
                      method = "ps",
                      ps_n.trees=1500,
                      ps_shrinkage=0.01,
                      ps_stop.method=c("es.max")
```

## **Index**

```
* datasets
    NSDUH_female, 8
    tMdat, 14
* models
    wgtmed, 15
* multivariate
    wgtmed, 15
* propensity
    twangMediation-package, 2
* score
    twangMediation-package, 2
bal.table.mediation, 2, 7, 11
calculate_effects, 3
check_missing, 4
data.frame, 16
desc.effects, 4, 5
desc.effects.mediation, 4, 5
dx.wts.mediation, 6, 13
formula, 16
gbm, 16, 17
NSDUH_female, 8
plot.mediation, 9
print.bal.table.mediation, 3, 10
print.mediation, 7, 11
ps, 7, 13, 17, 18
summary.mediation, 7, 12
swapTxCtrl, 13
tMdat, 14
twang, 17
twangMediation
        (twangMediation-package), 2
{\it twang Mediation-package}, 2
weighted_mean, 15
wgtmed, 3-5, 7, 8, 10-12, 14, 15
xgboost, 17
```