

Package ‘MultisiteMediation’

June 15, 2018

Version 0.0.2

Date 2018-06-14

Title Causal Mediation Analysis in Multisite Trials

Author Xu Qin, Guanglei Hong, Jonah Deutsch, and Edward Bein

Maintainer Xu Qin <xuqin@uchicago.edu>

Depends lme4, statmod, psych, MASS, ggplot2

Description We implement multisite causal mediation analysis using the methods proposed by Qin and Hong (2017) <doi:10.3102/1076998617694879> and Qin, Hong, Deutsch, and Bein (under review). It enables causal mediation analysis in multisite trials, in which individuals are assigned to a treatment or a control group at each site. It allows for estimation and hypothesis testing for not only the population average but also the between-site variance of direct and indirect effects. This strategy conveniently relaxes the assumption of no treatment-by-mediator interaction while greatly simplifying the outcome model specification without invoking strong distributional assumptions. This package also provides a function that can further incorporate a sample weight and a nonresponse weight for multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse, to enhance both the internal validity and external validity. Because the identification assumptions are not always warranted, the package also provides a weighting-based balance checking function for assessing the remaining overt bias, as well as a weighting-based sensitivity analysis function for further evaluating the potential bias related to omitted confounding or to propensity score model misspecification.

License GPL-2

RoxygenNote 6.0.1

URL <https://github.com/Xu-Qin/MultisiteMediation>

BugReports <https://github.com/Xu-Qin/MultisiteMediation/issues>

NeedsCompilation no

Repository CRAN

Date/Publication 2018-06-15 07:28:01 UTC

R topics documented:

balance	2
msmediate	4
msmediate.weights	5
sensitivity	6
sim	10
sim.weights	10
vartest.msmediate	11
vartest.msmediate.weights	12

Index	14
--------------	-----------

balance	<i>Balance checking for causal mediation analysis in multisite trials</i>
---------	---

Description

This function is used to check if, within a treatment group, the estimated nonresponse weight balances the distribution of the observed covariates between the respondents and the nonrespondents, or if the estimated RMPW weight balances the distribution of the observed covariates between those whose mediator takes value 1 and those whose mediator takes value 0.

Usage

```
balance(data, y, treatment, mediator, response, XR1, XR0, XM1, XM0, X, site)
```

Arguments

data	The data set for analysis.
y	The name of the outcome variable (string).
treatment	The name of the treatment variable (string).
mediator	The name of the mediator variable (string).
response	The name of the response variable (string), which is equal to 1 if the individual responded and 0 otherwise.
XR1	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
XR0	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.

XM1	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
XM0	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
X	A vector of variable names (string) of all the pretreatment covariates to be checked balance for.
site	The variable name for the site ID (string).

Value

A list of tables containing the balance checking results for the response before weighting (`$balance.R$balance1` under the treatment condition and `$balance.R$balance0` under the control condition) and after weighting (`$balance.R$balance1.adj` under the treatment condition and `$balance.R$balance0.adj` under the control condition); and the balance checking results for the mediator before weighting (`$balance.M$balance1` under the treatment condition and `$balance.M$balance0` under the control condition) and after weighting (`$balance.M$balance1.adj` under the treatment condition and `$balance.M$balance0.adj` under the control condition). It also contains a set of balance checking plots corresponding to the tables.

<code>balance.mean</code>	Population average of standardized bias. The standardized bias is calculated by dividing the unweighted (before weighting) or weighted (after weighting) mean difference between response or mediator levels in each covariate by the standard deviation of the covariate
<code>balance.sd</code>	Between-site standard deviation of standardized bias.
<code>balance.lower</code>	Lower bound of the 95% plausible value range of the site-specific standardized bias.
<code>balance.upper</code>	Upper bound of the 95% plausible value range of the site-specific standardized bias.

Author(s)

Xu Qin, Guanglei Hong, Jonah Deutsch, and Edward Bein

References

Qin, X., Hong, G., Deutsch, J., & Bein, E. (under review). A Multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse.

Examples

```
data(sim.weights)
```

```
balance(data = sim.weights, y = "y", treatment = "tr", mediator = "me", response = "R",
```

```
XR1 = c("x1", "x2", "x3"), XR0 = c("x1", "x2", "x3"), XM1 = c("x1", "x2", "x3"),
XM0 = c("x1", "x2", "x3"), X = c("x1", "x2", "x3"), site = "site")
```

msmediate

Causal mediation analysis in multisite trials

Description

This function is used to estimate both the population average and between-site variance of direct and indirect effects.

Usage

```
msmediate(data, y, treatment, mediator, X, site)
```

Arguments

data	The data set for analysis.
y	The name of the outcome variable (string).
treatment	The name of the treatment variable (string).
mediator	The name of the mediator variable (string).
X	A vector of variable names (string) of pretreatment covariates, which will be included in the propensity score model. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
site	The variable name for the site ID (string).

Value

A list contains the estimates of the between-site variance of direct effect, that of indirect effect, and the correlation between the direct and indirect effects across sites (`$Random_effects`), and the population average direct and indirect effect estimates along with their hypothesis testing results (`$Fixed_effects`).

Author(s)

Xu Qin and Guanglei Hong

References

Qin, X., & Hong, G (2017). A weighting method for assessing between-site heterogeneity in causal mediation mechanism. *Journal of Educational and Behavioral Statistics*. *Journal of Educational and Behavioral Statistics*, 42(3), 308-340. doi: [10.3102/1076998617694879](https://doi.org/10.3102/1076998617694879)

Examples

```
data(sim)

msmediate(data = sim, y = "y", treatment = "tr", mediator = "me", X = c("x1", "x2",
  "x3", "x4"), site = "site")
```

msmediate.weights	<i>Causal mediation analysis in multisite trials in the presence of complex sample and survey designs and non-random nonresponse</i>
-------------------	--

Description

This function is used to estimate both the population average and between-site variance of natural direct effect, natural indirect effect, pure indirect effect, and treatment-by-mediator interaction effect. It incorporates a sample weight to adjust for complex sample and survey designs and employs an estimated nonresponse weight to account for non-random nonresponse.

Usage

```
msmediate.weights(data, y, treatment, mediator, response, XR1, XR0, XM1, XM0, site,
  sample.weight)
```

Arguments

data	The data set for analysis.
y	The name of the outcome variable (string).
treatment	The name of the treatment variable (string).
mediator	The name of the mediator variable (string).
response	The name of the response variable (string), which is equal to 1 if the individual responded and 0 otherwise.
XR1	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
XR0	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
XM1	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.

<code>XM0</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>site</code>	The variable name for the site ID (string).
<code>sample.weight</code>	The variable name for the sample weight given by design (string).

Value

A list contains the estimates of the between-site variances of natural direct effect, natural indirect effect, pure indirect effect, and treatment-by-mediator interaction effect, and the correlations between the effects across sites (`$Random_effects`), and the population average effect estimates along with their hypothesis testing results (`$Fixed_effects`).

Author(s)

Xu Qin, Guanglei Hong, Jonah Deutsch, and Edward Bein

References

Qin, X., Hong, G., Deutsch, J., & Bein, E. (under review). A Multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse.

Examples

```
data(sim.weights)

msmediate.weights(data = sim.weights, y = "y", treatment = "tr", mediator = "me",
  response = "R", XR1 = c("x1", "x2", "x3"), XR0 = c("x1", "x2", "x3"), XM1 = c("x1",
    "x2", "x3"), XM0 = c("x1", "x2", "x3"), site = "site", sample.weight = "WD")
```

sensitivity

Sensitivity analysis for causal mediation analysis in multisite trials

Description

This function is used to calculate the effect size of the hidden bias associated with one or more omitted confounders or omitted random slopes of existing confounders, based on a small number of weighting-based sensitivity parameters.

Usage

```
sensitivity(data, y, treatment, mediator, response, XR1, XR0, XM1, XM0, omit.X = NULL,
  ran.omit.X = 1, site)
```

Arguments

<code>data</code>	The data set for analysis.
<code>y</code>	The name of the outcome variable (string).
<code>treatment</code>	The name of the treatment variable (string).
<code>mediator</code>	The name of the mediator variable (string).
<code>response</code>	The name of the response variable (string), which is equal to 1 if the individual responded and 0 otherwise.
<code>XR1</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the treatment condition in the original analysis. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>XR0</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the control condition in the original analysis. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>XM1</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the treatment condition in the original analysis. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>XM0</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the control condition in the original analysis. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>omit.X</code>	A vector of variable names (string) of omitted pretreatment confounders of the response-mediator, response-outcome, or mediator-outcome relationships.
<code>ran.omit.X</code>	A vector of variable names (string) of pretreatment covariates that have been included in the original analysis but whose random slopes are omitted from response models or mediator models in the original analysis. It takes value 1 if there are no omitted random slopes.
<code>site</code>	The variable name for the site ID (string).

Value

A list contains sensitivity parameters and the effect size of bias due to omission of pretreatment confounders ($\$Bias$ due to omission of confounders) or random slopes ($\$Bias$ due to omission of random slopes) for each causal parameter ($\$ITT$ for the population average ITT effect, $\$var.ITT$ for the between-site variance of ITT effect; $\$NIE$ for the population average natural indirect effect, $\$var.NIE$ for the between-site variance of natural indirect effect; $\$NDE$ for the population average natural direct effect, $\$var.NDE$ for the between-site variance of natural direct effect; $\$cov.NIE.NDE$ for the between-site covariance between natural indirect and direct effects; $\$PIE$ for the population average pure indirect effect; and $\$INT$ for the population average natural treatment-by-mediator interaction effect).

c1	Average conversion coefficient under the experimental condition.
sigma1.ITT	Average standard deviation of the ITT weight discrepancy in the experimental group at the same site. It is associated with the degree to which the omissions predict response under either treatment condition.
rho1.ITT	Average correlation between the ITT weight discrepancy and the outcome in the experimental group at the same site. It is related to the degree to which the omissions predict the outcome within levels of the response in the experimental group.
cov.sigma1.rho1.ITT	Covariance between site-specific sigma1.ITT and rho1.ITT.
c0	Average conversion coefficient under the control condition.
sigma0.ITT	Average standard deviation of the ITT weight discrepancy in the control group at the same site.
rho0.ITT	Average correlation between the ITT weight discrepancy and the outcome in the control group at the same site.
cov.sigma0.rho0.ITT	Covariance between site-specific sigma0.ITT and rho0.ITT.
bias.ITT	Effect size of bias of the population average ITT effect due to the omissions.
var.c.sigma.rho	Between-site variance of the effect size of bias for site-specific ITT.
cov.b.ITT.c.sigma.rho	Covariance between the effect size of site-specific ITT and the effect size of bias for site-specific ITT.
bias.var.ITT	Effect size of bias of the between-site variance of the ITT effect due to the omissions.
sigma.counter	Average standard deviation of the overall weight (product of the ITT weight and the RMPW weight) discrepancy in the experimental group at the same site. It is associated with the degree to which the omissions predict both response and mediator value assignment under either treatment condition.
rho.counter	Average correlation between the overall weight (product of the ITT weight and the RMPW weight) discrepancy and the outcome in the experimental group at the same site. It is related to the degree to which the omissions predict the outcome within levels of the response and mediator in the experimental group.
cov.sigma.rho.counter	Covariance between site-specific sigma.counter and rho.counter.
bias.NIE	Effect size of bias of the population average natural indirect effect due to the omissions.
var.c.sigma.rho.NIE	Between-site variance of the effect size of bias for site-specific NIE.
cov.b.NIE.c.sigma.rho	Covariance between the effect size of site-specific NIE and the effect size of bias for site-specific NIE.
bias.var.NIE	Effect size of bias of the between-site variance of the natural indirect effect due to the omissions.

<code>bias.NDE</code>	Effect size of bias of the population average natural direct effect due to the omissions.
<code>var.c.sigma.rho.NDE</code>	Between-site variance of the effect size of bias for site-specific NDE.
<code>cov.b.NDE.c.sigma.rho</code>	Covariance between the effect size of site-specific NDE and the effect size of bias for site-specific NDE.
<code>bias.var.NDE</code>	Effect size of bias of the between-site variance of the natural direct effect due to the omissions.
<code>bias.cov.NIE.NDE</code>	Effect size of bias of the between-site covariance between the natural direct and indirect effects due to the omissions.
<code>sigma.PIE</code>	Average standard deviation of the overall weight (product of the ITT weight and the RMPW weight) discrepancy in the control group at the same site.
<code>rho.PIE</code>	Average correlation between the overall weight (product of the ITT weight and the RMPW weight) discrepancy and the outcome in the control group at the same site.
<code>cov.sigma.rho.PIE</code>	Covariance between site-specific <code>sigma.PIE</code> and <code>rho.PIE</code> .
<code>bias.PIE</code>	Effect size of bias of the population average pure indirect effect due to the omissions.
<code>bias.INT</code>	Effect size of bias of the population average natural treatment-by-mediator interaction effect due to the omissions.

Author(s)

Xu Qin, Guanglei Hong, Jonah Deutsch, and Edward Bein

References

Qin, X., Hong, G., Deutsch, J., & Bein, E. (under review). A Multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse.

Examples

```
data(sim.weights)

sensitivity(data = sim.weights, y = "y", treatment = "tr", mediator = "me",
  response = "R", XR1 = "x2", XR0 = "x2", XM1 = c("x1", "x2"), XM0 = "x2",
  omit.X = c("x1", "x3"), ran.omit.X = "x2", site = "site")
```

sim	<i>A simulated example data</i>
-----	---------------------------------

Description

This simulated data list is for demonstration.

Value

A list containing

site	Site ID
y	Outcome
tr	Treatment
me	Mediator
x1	Pretreatment covariate
x2	Pretreatment covariate
x3	Pretreatment covariate
x4	Pretreatment covariate

sim.weights	<i>A simulated example data</i>
-------------	---------------------------------

Description

This simulated data list is for demonstration.

Value

A list containing

site	Site ID
y	Outcome
tr	Treatment
me	Mediator
x1	Pretreatment covariate
x2	Pretreatment covariate
x3	Pretreatment covariate
R	Response indicator
WD	Sample weight

vartest.msmediate *Variance testing for multisite causal mediation analysis*

Description

This function performs hypothesis testing for the between-site variance of direct effect and that of indirect effect, besides providing the same output as given by the function `msmediate()`.

Usage

```
vartest.msmediate(data, y, treatment, mediator, X, site, npermute = 200)
```

Arguments

<code>data</code>	The data set for analysis.
<code>y</code>	The name of the outcome variable (string).
<code>treatment</code>	The name of the treatment variable (string).
<code>mediator</code>	The name of the mediator variable (string).
<code>X</code>	A vector of variable names (string) of pretreatment covariates, which will be included in the propensity score model. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>site</code>	The variable name for the site ID (string).
<code>npermute</code>	The number of permutations for the permutation test. The default value is 200. It may take a long time, depending on the sample size and the length of X.

Value

A list contains the hypothesis testing results of the between-site variance of the causal effects, besides the same output as given by the function `msmediate()`.

Author(s)

Xu Qin and Guanglei Hong

References

Qin, X., & Hong, G (2017). A weighting method for assessing between-site heterogeneity in causal mediation mechanism. *Journal of Educational and Behavioral Statistics*. *Journal of Educational and Behavioral Statistics*. *Journal of Educational and Behavioral Statistics*, 42(3), 308-340. doi: [10.3102/1076998617694879](https://doi.org/10.3102/1076998617694879)

Examples

```
data(sim)

vartest.msmediate(data = sim, y = "y", treatment = "tr", mediator = "me", X = c("x1",
  "x2", "x3", "x4"), site = "site", npermute = 2)
```

```
vartest.msmediate.weights
```

Variance testing for multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse

Description

This function performs hypothesis testing for the between-site variances of natural direct effect, natural indirect effect, pure indirect effect, and treatment-by-mediator interaction effect in the presence of complex sample and survey designs and non-random nonresponse, besides providing the same output as given by the function `msmediate.weights()`.

Usage

```
vartest.msmediate.weights(data, y, treatment, mediator, response, XR1, XR0, XM1,
  XM0, site, sample.weight, npermute = 200)
```

Arguments

<code>data</code>	The data set for analysis.
<code>y</code>	The name of the outcome variable (string).
<code>treatment</code>	The name of the treatment variable (string).
<code>mediator</code>	The name of the mediator variable (string).
<code>response</code>	The name of the response variable (string), which is equal to 1 if the individual responded and 0 otherwise.
<code>XR1</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
<code>XR0</code>	A vector of variable names (string) of pretreatment covariates in the propensity score model for the response under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.

XM1	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the treatment condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
XM0	A vector of variable names (string) of pretreatment covariates in the propensity score model for the mediator under the control condition. The function will treat covariates with fewer than 10 unique values as categorical variables. For now, the multilevel propensity score model only allows for one random intercept.
site	The variable name for the site ID (string).
sample.weight	The variable name for the sample weight given by design (string).
npermute	The number of permutations for the permutation test. The default value is 200. It may take a long time, depending on the sample size and the length of X.

Value

A list contains the hypothesis testing results of the between-site variance of the causal effects, besides the same output as given by the function `msmediate()`.

Author(s)

Xu Qin, Guanglei Hong, Jonah Deutsch, and Edward Bein

References

Qin, X., Hong, G., Deutsch, J., & Bein, E. (under review). A Multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse.

Examples

```
data(sim.weights)

vartest.msmediate.weights(data = sim.weights, y = "y", treatment = "tr",
  mediator = "me", response = "R", XR1 = c("x1", "x2", "x3"), XR0 = c("x1",
    "x2", "x3"), XM1 = c("x1", "x2", "x3"), XM0 = c("x1", "x2", "x3"),
  site = "site", sample.weight = "WD", npermute = 2)
```

Index

balance, [2](#)

msmediate, [4](#)

msmediate.weights, [5](#)

sensitivity, [6](#)

sim, [10](#)

sim.weights, [10](#)

vartest.msmediate, [11](#)

vartest.msmediate.weights, [12](#)