

Package ‘mma’

February 24, 2016

Type Package

Title Multiple Mediation Analysis

Version 2.0-3

Date 2016-2-27

Author Qingzhao Yu

Maintainer Qingzhao Yu <qyu@lsuhsc.edu>

Depends R (>= 2.14.1), gbm

Description Used for general multiple mediation analysis.

The analysis method is described in Yu et al. 2014, “General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival”, published on Journal of Biometrics & Biostatistics, 5(2):189.

License GPL (>= 2)

URL <http://www.r-project.org>,
http://publichealth.lsuhs.edu/Faculty_pages/qyu/index.html

NeedsCompilation no

Repository CRAN

Date/Publication 2016-02-24 16:59:20

R topics documented:

mma-package	2
boot.med.binx	3
boot.med.contx	6
data.org	8
med.binx	11
med.contx	13
mma	15
plot.mma	18
print.mma	20
summary.med_iden	21
summary.mma	22
weight_behavior	23

mma-package

Mediation Analysis Package

Description

This package is used to identify mediators and for general mediation analysis. Mediation effect refers to the effect conveyed by intervening variables to an observed relationship between an exposure and a response variable (outcome). In this package, the exposure is called the predictor, the intervening variables are called mediators. The mediation effects include the total effect, direct effect, and indirect effect, which are defined and the statistical inferences described in Yu et al.(2014).

Details

Package: mma
Type: Package
Version: 1.0
Date: 2015-02-18
License: GPL (>= 2)

"[data.org](#)" is used to identify potential mediators. "[med.binx](#)", "[boot.med.binx](#)", "[med.contx](#)", and "[boot.med.contx](#)" are used for statistical inferences on the mediation effects when the predictor is binary or continuous according to the name of the function. "[mma](#)" is a combined function that automatically identify potential mediators, based on which to make statistical inference on the mediation effects.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

Maintainer: Qingzhao Yu <qyu@lsuhsc.edu>

References

Baron, R.M., and Kenny, D.A. (1986). "The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations," *J. Pers Soc Psychol*, 51(6), 1173-1182.

J.H. Friedman, T. Hastie, R. Tibshirani (2000). "Additive Logistic Regression: a Statistical View of Boosting," *Annals of Statistics* 28(2):337-374.

J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29(5):1189-1232.

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

Examples

```

data("weight_behavior")
#binary x
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=FALSE,nu=0.001,
  D=3,distn="bernoulli",family1=binomial(link = "logit"),n2=2)
temp.b.b.mart<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=TRUE,nu=0.05,
  D=3,family1=binomial(link = "logit"),n2=5)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
temp.b.c.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=8:10),biny=FALSE,
  predref="M",alpha=0.4,alpha2=0.4,n=2,seed=1,mart=FALSE,nu=0.05,D=3,
  distn="gaussian",family1=gaussian(link = "identity"),n2=2)
temp.b.c.mart<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=8:10, j2=12:13),biny=FALSE,
  predref="M",alpha=0.4,alpha2=0.4, margin=1, n=2,seed=1,mart=TRUE,
  nu=0.05,distn="gaussian",family1=gaussian(link = "identity"),n2=2)

```

boot.med.binx

Inferences on Mediation Effects with Binary Predictor

Description

To make inferences on the mediation effects when the predictor is binary.

Usage

```

boot.med.binx(x, y, dirx, contm = NULL, catm = NULL, jointm = NULL,
  n = 20, seed = sample(1:1000, 1), n2 = 50, mart = F, nu = 0.001,
  D = 3, distn = "bernoulli", family1 = binomial("logit"),
  weight=rep(1,length(y)))

```

Arguments

x	a data frame contains the predictor, all mediators and covariates.
y	the vector of outcome variable.
dirx	the column number of x that locates the predictor.
contm	the column numbers of x that locate the potential continuous mediators.

catm	the column numbers of x that locate the potential binary or categorical mediators.
jointm	a list where the first item is the number of groups of joint mediators to be considered, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
n	the time of resampling when calculating the indirect effects, see Yu et al. (2014). The default value is n=20.
seed	set seed to make the calculation repeatable. The default value of seed is sampled by <code>sample(1:1000,1)</code> .
n2	the number of times of bootstrap resampling. The default value is n2=50.
mart	if TRUE, Multiple Additive Regression Trees (MART) will be used to fit the final full model in estimating the outcome. The default value of mart is FALSE, in which case a generalized linear model will be used to fit the final full model.
nu	set the parameter "interaction.depth" in gbm function if MART is to be used, by default, nu=0.001. See also <code>help(gbm.fit)</code> .
D	set the parameter "shrinkage" in gbm function if MART is to be used, by default, D=3. See also <code>help(gbm.fit)</code> .
distn	the assumed distribution of the outcome if MART is used for final full model. The default value of distn is "bernoulli". If y is continuous, may use <code>distn="gaussian"</code> .
family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component if generalized linear model is used as the final full model. The default value of family1 is <code>binomial(link="logit")</code> . If y is continuous, family1 can be defined as <code>gaussian(link="identity")</code> .
weight	the weight for observations.

Details

The potential mediators are not tested in this function. `data.org` should be used for the tests and the resulted list can be used directly to define the parameters in this function. `boot.med.binx` considers all variables in x as mediators or covariates in the final model and all variables identified by `contm`, `catm`, or `jointm` as mediators.

Value

Returns an `mma` object.

estimation	list the estimation of ie (indirect effect), te (total effect), and de (direct effect from the predictor) separately.
bootsresults	a list where the first item, ie, is a matrix of n2 rows where each column gives the estimated indirect effect from the corresponding mediator (identified by the column name) from the n2 bootstrap samples; the second item, te, is a vector of estimated total effects from the bootstrap sample; and the 3rd item, de, is a vector of estimated direct effect of the predictor from the bootstrap sample.
model	a list where the first item, mart, is T if MART is fitted for the final model; the second item, model, is the fitted final full model where y is the outcome and all predictor, covariates, and mediators are the explanatory variables; and the third

item, best.iter is the number of best iterations if MART is used to fit the final model.

data a list that contain the all the used data: x=x,y=y,dirx=dirx,contm=contm,catm=catm,jointm=jointm, binpred=T

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"[med.binx](#)" to estimate the mediation effects, "[med.contx](#)", and "[boot.med.contx](#)" for mediation analysis when the predictor is continuous.

Examples

```
data("weight_behavior")
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.bin<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4)
temp1<-boot.med.binx(x=data.bin$x,y=data.bin$y,dirx=data.bin$dirx,
  contm=data.bin$contm,catm=data.bin$catm,n=2,n2=4,seed=1)
temp2<-boot.med.binx(x=data.bin$x,y=data.bin$y,dirx=data.bin$dirx,contm=data.bin$contm,
  catm=data.bin$catm,n=2,n2=4,nu=0.05,seed=1,mart=TRUE)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
data.cont<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,biny=FALSE,
  family1=gaussian(link="identity"),predref="M",alpha=0.4,alpha2=0.4)
temp3<-boot.med.binx(data.cont$x,y=data.cont$y,dirx=data.cont$dirx,
  contm=data.cont$contm,catm=data.cont$catm,n=2,n2=4,seed=1,
  family1=gaussian(link="identity"))
temp4<-boot.med.binx(data.cont$x,y=data.cont$y,dirx=data.cont$dirx,contm=data.cont$contm,
  catm=data.cont$catm,n=2,n2=4,nu=0.05,seed=1, mart=TRUE,distn="gaussian")
```

Description

To make inferences on the mediation effects when the predictor is continuous.

Usage

```
boot.med.contx(x, y, dirx, binm = NULL, contm = NULL, catm = NULL,
  jointm = NULL, margin = 1, n = 20, seed = sample(1:1000, 1),
  mart = F, nu = 0.001, D = 3, distn = "gaussian",
  family1 = gaussian(link = "identity"), n2 = 50, weight=rep(1,length(y)))
```

Arguments

x	a data frame contains the predictor, all mediators and covariates.
y	the vector of outcome variable.
dirx	the column number of x that locates the predictor.
binm	the column number of x that locates the binary mediators.
contm	the column numbers of x that locate the potential continuous mediators.
catm	categorical mediators should be binarized and be presented as a list, where the first item is the number of categorical variables and the following items are the column numbers of each binarized categorical variable in x. data.org organize the categorical mediators in this format after they pass the mediator tests.
jointm	a list where the first item is the number of groups of joint mediators to be considered, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
margin	the change in predictor when calculating the mediation effects, see Yu et al. (2014).
n	the time of resampling in calculating the indirect effects, default is n=20, see Yu et al. (2014).
seed	set seed to make the calculation repeatable. The default value of seed is from sample(1:1000,1).
mart	if TURE, Multiple Additive Regression Trees (MART) will be used to fit the final full model in estimating the outcome. The default value of mart is FALSE, in which case, a generalized linear model will be used to fit the final full model.
nu	set the parameter "interaction.depth" in gbm function if MART is to be used, by default, nu=0.001. See also help(gbm.fit).
D	set the parameter "shrinkage" in gbm function if MART is to be used, by default, D=3. See also help(gbm.fit).
distn	the assumed distribution of the outcome if MART is used for final full model. The default value of distn is "gaussian". If y is binary, may use distn="bernoulli".

family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component if generalized linear model is used as the final full model. The default value of family1 is gaussian(link="identity"). If y is binary, family1 can be defined as binomial(link = "logit").
n2	the number of times of bootstrap resampling. The default value is n2=50.
weight	the weight for observations.

Details

The mediators are not tested in this function. `data.org` should be used for the tests and the resulted list can be used directly to define the parameters in this function. `boot.med.contx` considers all variables in `x` as mediators or covariates in the final model and all variables identified by `contm`, `binm`, `catm`, or `jointm` as mediators.

Value

Returns an `mma` object.

estimation	list the estimation of <code>ie</code> (indirect effect), <code>te</code> (total effect), and <code>de</code> (direct effect from the predictor) separately.
bootsresults	a list where the first item, <code>ie</code> , is a matrix of <code>n2</code> rows where each column gives the estimated indirect effect from the corresponding mediator (identified by the column name) from the <code>n2</code> bootstrap samples; the second item, <code>te</code> , is a vector of estimated total effects from the bootstrap sample; and the 3rd item, <code>de</code> , is a vector of estimated direct effect of the predictor from the bootstrap sample.
model	a list where the first item, <code>mart</code> , is T if MART is fitted for the final model; the second item, <code>model</code> , is the fitted final full model where <code>y</code> is the outcome and all predictor, covariates, and mediators are the explanatory variables; and the third item, <code>best.iter</code> is the number of best iterations if MART is used to fit the final model.
data	a list that contain all the used data: <code>x=x,y=y,dirx=dirx,binm=binm,contm=contm,catm=catm,jointm=jointm, binpred=F</code>

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"[med.contx](#)" to estimate mediation effects, "[med.binx](#)" and "[boot.med.binx](#)" for mediation analysis when the predictor is binary.

Examples

```

data("weight_behavior")
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.contx<-data.org(x,y,pred=1,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,binpred=FALSE,alpha=0.4,alpha2=0.4)
temp5<-boot.med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,
  binm=data.contx$binm, contm=data.contx$contm,catm=data.contx$catm,
  seed=1,n=1,family1=binomial(link = "logit"),n2=2)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
data.contx<-data.org(x,y,pred=1,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,binpred=FALSE,
  family1=gaussian(link="identity"),alpha=0.4,alpha2=0.4)
temp7<-boot.med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,
  binm=data.contx$binm, contm=data.contx$contm,catm=data.contx$catm,
  seed=1,n=1,n2=2)

```

data.org

Data Organization Considering Joint Mediators

Description

Do a preliminary data analysis to identify potential mediators and covariates. Each variable listed in jointm is forced in the final estimation model as a mediator. Also organize the data into a format that can be directly used for the mediation analysis functions.

Usage

```

data.org(x, y, pred, contmed = NULL, binmed = NULL,
  binref = rep(1,length(binmed)), catmed = NULL,
  catref = rep(1,length(catmed)), jointm = NULL, biny = T,
  family1 = binomial(link = "logit"), binpred = T, predref = 1,
  alpha = 0.1, alpha2 = 0.1)

```

Arguments

x	a data frame contains the predictor, all potential mediators and covariates
y	the vector of outcome variable.
pred	the column number that locates the predictor in x.
contmed	a vector of column numbers that locate the potential continuous mediators in x.
binmed	a vector of column numbers that locate the potential binary mediators in x.
binref	the defined reference groups of the binary potential mediators in binmed.

catmed	a vector of column numbers that locate the potential categorical mediators in x.
catref	the defined reference groups of the categorical potential mediators in catmed.
jointm	a list that identifies the mediators that need to be considered jointly, where the first item indicates the number of groups of mediators to be considered jointly, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
biny	if TRUE, the outcome variable is binary. The default value of biny is TRUE.
family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component in generalized linear model. The default value of family1 is binomial(link = "logit"). If y is continuous, family1 can be defined as gaussian(link="identity").
binpred	if TRUE, the predictor is binary. The default value of binpred is TRUE.
predref	if binpred is TRUE, identify the reference group of the binary predictor.
alpha	the significance level at which to test if the potential mediators (identified by contmed, binmed, and catmed) can be used as a covariate or mediator in estimating y when all variables in x are included in the model. The default value is alpha=0.1
alpha2	the significant level at which to test if a potential mediator is related with the predictor. The default value is alpha2=0.1.

Value

data.org returns a list with the organized data and identifiers of the potential mediators in the organized data set.

x	the organized data frame that include all potential mediators and covariates that should be used to estimate the outcome.
dirx	the column number of x that locates the predictor.
contm	the column numbers of x that locate the potential continuous mediators.
binm	when the predictor is continuous, binm gives the column numbers of x that locate the potential binary mediators.
catm	when the predictor is binary, catm gives the column numbers of x that locate the potential binary or categorical mediators; when the predictor is continuous, catm gives a list where the first item is the number of potential categorical mediators, and the following items give the column numbers of each binarized categorical mediator in x.
jointm	a list where the first item is the number of groups of joint mediators, and each of the following items identifies the column numbers of the mediators in the newly organized x for each group of joint mediators.
fullmodel	the full linear model fitted with all potential mediators and covariates.
rela	p-values of tests on the relationship between the predictor and each potential mediator.

Note

All other variables in *x* but not identified by *contmed*, *binmed*, or *catmed* are forced in the final model as covariates. Compared with *data.org*, joint mediators are considered in this function. Every variable in the *jointm* should be listed in *contmed*, *binmed*, or *catmed*, and these variables are forced to be included as mediators for further mediation analysis. A variables can be included in more than one groups of joint mediators in *jointm*.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Baron, R.M., and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers Soc Psychol*, 51(6), 1173-1182.

See Also

"[data.org](#)" that does not consider joint mediators, which can be added freely in the mediation analysis functions later.

Examples

```
data("weight_behavior")
#binary x
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.b.b.2<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=c(6,8,10)),predref="M",
  alpha=0.4,alpha2=0.4)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
data.b.c.2<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=8:10),biny=FALSE,
  family1=gaussian(link="identity"),predref="M",alpha=0.4,alpha2=0.4)

#continuous x
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.c.b.2<-data.org(x,y,pred=1,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=2,j1=8:10,j2=c(6,8)),
  binpred=FALSE,alpha=0.4,alpha2=0.4)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
```

```
data.c.c.2<-data.org(x,y,pred=1,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=2,j1=8:10,j2=c(6,8)),
  binpred=FALSE,family1=gaussian(link="identity"),alpha=0.4,alpha2=0.4)
```

med.binx

*Mediation Analysis with Binary Predictor***Description**

To estimate the mediation effects when the predictor is binary.

Usage

```
med.binx(x, y, dirx, contm = NULL, catm = NULL, jointm = NULL,
  allm = c(contm, catm), n = 20, seed = sample(1:1000, 1), mart = F,
  nu = 0.001, D = 3, distn = "bernoulli", family1 = binomial("logit"))
```

Arguments

x	a data frame contains the predictor, all mediators and covariates.
y	the vector of outcome variable.
dirx	the column number of x that locates the predictor.
contm	the column numbers of x that locate the potential continuous mediators.
catm	the column numbers of x that locate the potential binary or categorical mediators.
jointm	a list where the first item is the number of groups of joint mediators to be considered, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
allm	the column numbers of all mediators. The default value of allm is c(contm,catm).
n	the time of resampling when calculating the indirect effects, see Yu et al. (2014). The default value is n=20.
seed	set seed to make the calculation repeatable. The default value of seed is sampled by sample(1:1000,1).
mart	if TURE, Multiple Additive Regression Trees (MART) will be used to fit the final full model in estimating the outcome. The default value of mart is FALSE, in which case a generalized linear model will be used to fit the final full model.
nu	set the parameter "interaction.depth" in gbm function if MART is to be used, by default, nu=0.001. See also help(gbm.fit).
D	set the parameter "shrinkage" in gbm function if MART is to be used, by default, D=3. See also help(gbm.fit).
distn	the assumed distribution of the outcome if MART is used for final full model. The default value of distn is "bernoulli". If y is continuous, may use distn="gaussian".
family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component if generalized linear model is used as the final full model. The default value of family1 is binomial(link = "logit"). If y is continuous, family1 can be defined as gaussian(link="identity").

Details

The mediators are not tested in this function. `data.org` should be used for the tests and the resulted list can be used directly to define the parameters in this function. `med.binx` considers all variables in `x` as mediators or covariates in the final model and all variables identified by `contm`, `catm`, `allm`, or `jointm` as mediators.

Value

<code>denm</code>	a matrix where each column gives the estimated direct effect not from the corresponding mediator (identified by the column name) and each row corresponding to the results from one resampling.
<code>ie</code>	a matrix where each column gives the estimated indirect effect from the corresponding mediator (identified by the column name) and each row corresponding to the results from one resampling.
<code>te</code>	a vector of the estimated total effect from each resampling.
<code>model</code>	the fitted final full model where <code>y</code> is the outcome and all predictor, covariates, and mediators are the explanatory variables.
<code>best.iter</code>	the number of best iterations if MART is used to fit the final model, is <code>NULL</code> if the final model is a generalized linear model.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

- J.H. Friedman, T. Hastie, R. Tibshirani (2000). "Additive Logistic Regression: a Statistical View of Boosting," *Annals of Statistics* 28(2):337-374.
- J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29(5):1189-1232.
- Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"[boot.med.binx](#)" to make inferences on the estimated mediation effects using bootstrap method, "[med.contx](#)", and "[boot.med.contx](#)" for mediation analysis when the predictor is continuous.

Examples

```
data("weight_behavior")
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.bin<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4)
temp1<-med.binx(x=data.bin$x,y,dirx=data.bin$dirx,
```

```

contm=data.bin$contm,catm=data.bin$catm,n=2,seed=1)
temp2<-med.binx(x=data.bin$x,y,dirx=data.bin$dirx,
  contm=data.bin$contm,catm=data.bin$catm,n=2,seed=1,mart=TRUE)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
data.cont<-data.org(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,biny=FALSE,
  family1=gaussian(link="identity"),predref="M",alpha=0.4,alpha2=0.4)
temp3<-med.binx(data.cont$x,y,dirx=data.cont$dirx,contm=data.cont$contm,
  catm=data.cont$catm,n=2,seed=1,family1=gaussian(link="identity"))
temp4<-med.binx(data.cont$x,y,dirx=data.cont$dirx,contm=data.cont$contm,
  catm=data.cont$catm,n=2,seed=1,mart=TRUE,distn="gaussian")

```

med.contx

Mediation Analysis with Continuous Predictor

Description

To estimate the mediation effects when the predictor is continuous.

Usage

```

med.contx(x, y, dirx, binm = NULL, contm = NULL, catm = NULL,
  jointm = NULL, margin = 1, n = 20, seed = sample(1:1000, 1),
  mart = F, nu = 0.001, D = 3, distn = "gaussian",
  family1 = gaussian(link = "identity"))

```

Arguments

x	a data frame contains the predictor, all mediators and covariates.
y	the vector of outcome variable.
dirx	the column number of x that locates the predictor.
binm	the column number of x that locates the binary mediators.
contm	the column numbers of x that locate the potential continuous mediators.
catm	categorical mediators should be binarized and be presented as a list, where the first item is the number of categorical variables and the following items are the column numbers of each binarized categorical variable in x. data.org organizes the categorical mediators in this format after they pass the mediator tests.
jointm	a list where the first item is the number of groups of joint mediators to be considered, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
margin	the change in predictor when calculating the mediation effects, see Yu et al. (2014).

n	the time of resampling in calculating the indirect effects, default is n=20, see Yu et al. (2014).
seed	set seed to make the calculation repeatable. The default value of seed is from sample(1:1000,1).
mart	if TRUE, Multiple Additive Regression Trees (MART) will be used to fit the final full model in estimating the outcome. The default value of mart is FALSE, in which case, a generalized linear model will be used to fit the final full model.
nu	set the parameter "interaction.depth" in gbm function if MART is to be used, by default, nu=0.001. See also help(gbm.fit).
D	set the parameter "shrinkage" in gbm function if MART is to be used, by default, D=3. See also help(gbm.fit).
distn	the assumed distribution of the outcome if MART is used for final full model. The default value of distn is "gaussian". If y is binary, may use distn="bernoulli".
family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component if generalized linear model is used as the final full model. The default value of family1 is gaussian(link="identity"). If y is binary, family1 can be defined as binomial(link = "logit").

Details

The mediators are not tested in this function. data.org should be used for the tests and the resulted list can be used directly to define the parameters in this function. med.contx considers all variables in x as mediators or covariates in the final model and all variables identified by contm, binm, catm, or jointm as mediators.

Value

denm	a matrix where each column gives the estimated direct effect not from the corresponding mediator (identified by the column name), see Yu et al. (2014) for the definition, and each row corresponding to the results from one resampling.
ie	a matrix where each column gives the estimated indirect effect from the corresponding mediator (identified by the column name) and each row corresponding to the results from one resampling.
te	a vector of the estimated total effect from each resampling.
model	a list, where the first item, mart, is TRUE if a MART is fitted as the final model; the second item, full.model, is the fitted final full model where y is the outcome and all predictor, covariates, and mediators are the explanatory variables; and the third item, best.iter, is the number of best iterations if MART is used to fit the final model, is NULL if the final model is a generalized linear model.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

- J.H. Friedman, T. Hastie, R. Tibshirani (2000). "Additive Logistic Regression: a Statistical View of Boosting," *Annals of Statistics* 28(2):337-374.
- J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29(5):1189-1232.
- Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"[boot.med.contx](#)" to make inferences on the estimated mediation effects using bootstrap method, "[med.binx](#)" and "[boot.med.binx](#)" for mediation analysis when the predictor is binary.

Examples

```
data("weight_behavior")
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
data.contx<-data.org(x,y,pred=1,contmed=c(5,8:10),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,binpred=FALSE,alpha=0.4,alpha2=0.4)
temp5<-med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,binm=data.contx$binm,
  contm=data.contx$contm,catm=data.contx$catm,seed=1,n=2,
  family1=binomial(link = "logit"))
temp6<-med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,binm=data.contx$binm,
  contm=data.contx$contm,catm=data.contx$catm,seed=1,n=2,mart=TRUE,nu=0.05,
  distn="bernoulli")

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
data.contx<-data.org(x,y,pred=1,contmed=c(12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,binpred=FALSE,
  family1=gaussian(link="identity"),alpha=0.4,alpha2=0.4)
temp7<-med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,binm=data.contx$binm,
  contm=data.contx$contm,catm=data.contx$catm,seed=1,n=2)
temp8<-med.contx(x=data.contx$x,y=data.contx$y,dirx=data.contx$dirx,binm=data.contx$binm,
  contm=data.contx$contm,catm=data.contx$catm,seed=1,n=2,mart=TRUE,nu=0.05)
```

Description

Test for mediators and do statistical inferences on the identified mediation effects.

Usage

```
mma(x, y, pred, contmed = NULL, binmed = NULL,
    binref = rep(1,length(binmed)), catmed = NULL,
    catref = rep(1,length(catmed)), jointm = NULL, biny = T, binpred = T,
    predref = 1, alpha = 0.1, alpha2 = 0.1, margin = 1, n = 20,
    seed = sample(1:1000, 1), mart = F, nu = 0.001, D = 3,
    distn = "bernoulli", family1 = binomial(link = "logit"), n2 = 50,
    weight=rep(1,length(y)))
```

Arguments

x	a data frame contains the predictor, all potential mediators and covariates
y	the vector of outcome variable.
pred	the column number that locates the predictor in x.
contmed	a vector of column numbers that locate the potential continuous mediators in x.
binmed	a vector of column numbers that locate the potential binary mediators in x.
binref	the defined reference groups of the binary potential mediators in binmed.
catmed	a vector of column numbers that locate the potential categorical mediators in x.
catref	the defined reference groups of the categorical potential mediators in catmed.
jointm	a list that identifies the mediators that need to be considered jointly, where the first item indicates the number of groups of mediators to be considered jointly, and each of the following items identifies the column numbers of the mediators in x for each group of joint mediators.
biny	if TRUE, the outcome variable is binary. The default value of biny is TRUE.
binpred	if TRUE, the predictor is binary. The default value of binpred is TRUE.
predref	if binpred is TRUE, identify the reference group of the binary predictor.
alpha	the significance level at which to test if the potential mediators (identified by contmed, binmed, and catmed) can be used as a covariate or mediator in estimating y when all variables in x are included in the model. The default value is alpha=0.1
alpha2	the significant level at which to test if a potential mediator is related with the predictor. The default value is alpha2=0.1.
margin	if binpred is FALSE, define the change in predictor when calculating the mediation effects, see Yu et al. (2014).
n	the time of resampling in calculating the indirect effects, default is n=20, see Yu et al. (2014).
seed	set seed to make the calculation repeatable. The default value of seed is from sample(1:1000,1).
mart	if TRUE, Multiple Additive Regression Trees (MART) will be used to fit the final full model in estimating the outcome. The default value of mart is FALSE, in which case, a generalized linear model will be used to fit the final full model.
nu	set the parameter "interaction.depth" in gbm function if MART is to be used, by default, nu=0.001. See also help(gbm.fit).

D	set the parameter "shrinkage" in gbm function if MART is to be used, by default, D=3. See also help(gbm.fit).
distn	the assumed distribution of the outcome if MART is used for final full model. The default value of distn is "bernoulli". If y is continuous, may use distn="gaussian".
family1	define the conditional distribution of y given x, and the linkage function that links the mean of y with the system component if generalized linear model is used as the final full model. The default value of family1 is binomial(link = "logit"). If y is binary, family1 can be defined as gaussian(link="identity").
n2	the number of times of bootstrap resampling. The default value is n2=50.
weight	the weight for each observation.

Details

mma first tests if the potential mediators defined by binm, contm, and catm should be treated as mediators or covariates (if none, the variable will be deleted from further analysis). All variables identified by jointm are treated as mediators. All other variables in x that are not tested are treated as covariates. Then mma does the mediation effects estimation and inference on the selected variables.

Value

Returns an mma object.

estimation	list the estimation of ie (indirect effect), te (total effect), and de (direct effect from the predictor) separately.
bootsresults	a list where the first item, ie, is a matrix of n2 rows where each column gives the estimated indirect effect from the corresponding mediator (identified by the column name) from the n2 bootstrap samples; the second item, te, is a vector of estimated total effects from the bootstrap sample; and the 3rd item, de, is a vector of estimated direct effect of the predictor from the bootstrap sample.
model	a list where the first item, mart, is T if MART is fitted for the final model; the second item, model, is the fitted final full model where y is the outcome and all predictor, covariates, and mediators are the explanatory variables; and the third item, best.iter is the number of best iterations if MART is used to fit the final model.
data	a list that contain all the used data

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

- Baron, R.M., and Kenny, D.A. (1986). "The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations," *J. Pers Soc Psychol*, 51(6), 1173-1182.
- J.H. Friedman, T. Hastie, R. Tibshirani (2000). "Additive Logistic Regression: a Statistical View of Boosting," *Annals of Statistics* 28(2):337-374.

J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29(5):1189-1232.

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"[data.org](#)" is for mediator tests, and "[med.binx](#)", "[boot.med.binx](#)", "[med.contx](#)", and "[boot.med.contx](#)" for mediation analysis and inferences.

Examples

```
data("weight_behavior")
#binary x
#binary y
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=FALSE,nu=0.001,
  D=3,n2=2)
temp.b.b.mart<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=TRUE,nu=0.05,
  D=3,n2=5)

#continuous y
x=weight_behavior[,2:14]
y=weight_behavior[,1]
temp.b.c.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=8:10),biny=FALSE,
  predref="M",alpha=0.4,alpha2=0.4,n=2,seed=1,mart=FALSE,nu=0.05,D=3,
  distn="gaussian",family1=gaussian(link = "identity"),n2=2)
temp.b.c.mart<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,jointm=list(n=1,j1=8:10, j2=12:13),biny=FALSE,
  predref="M",alpha=0.4,alpha2=0.4, margin=1, n=2,seed=1,mart=TRUE,
  nu=0.05,distn="gaussian",family1=gaussian(link = "identity"),n2=2)
```

plot.mma

Relative effects plot of the fitted mma object

Description

Plot the marginal effect of the selected variable on the outcome, and the marginal effect of the predictor on the selected variable.

Usage

```
## S3 method for class 'mma'
plot(x, ..., vari, xlim=range(data$x[, vari], na.rm=T))
```

Arguments

x	a mma object created initially call to mma, boot.med.binx, or boot.med.contx.
vari	an indices or the name of the variable to plot.
xlim	the range of the variable to be plotted.
...	other arguments passed to the plot function.

Details

plot.mma plots the marginal effect of the selected variable on the outcome, and the marginal effect of the predictor on the selected variable. If the predictor is binary, draw a histogram or boxplot of the marginal density of the variable at each different value of the predictor.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

"mma", "boot.med.binx", "boot.med.contx"

Examples

```
data("weight_behavior")
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),binref=c(1,1),
  catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4, jointm=NULL,margin=1, n=2,
  seed=1,mart=FALSE,nu=0.001,D=3,distn="bernoulli",
  family1=binomial(link = "logit"),n2=2)
plot(temp.b.b.glm,vari="exercises",xlim=c(0,50))
```

```
print.mma          Print a mma object
```

Description

Print the estimation of mediation effects from an mma object.

Usage

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

Arguments

x a mma object created initially call to mma, boot.med.binx, or boot.med.contx.
 ... other arguments passed to the print function.

Value

Return a LIST

ie estimation of the indirect effects from the mma object.
 te estimation of the total effect from the mma object.
 de estimation of the direct effect from the mma object.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

["mma"](#), ["boot.med.binx"](#), ["boot.med.contx"](#)

Examples

```
data("weight_behavior")
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=FALSE,nu=0.001,
  D=3,distn="bernoulli",family1=binomial(link = "logit"),n2=2)
print(temp.b.b.glm)
```

summary.med_iden *Summary method for class "med_iden".*

Description

Compute the estimations, standard deviations and confidence intervals of the mediation effects.

Usage

```
## S3 method for class 'med_iden'
summary(object,...)
## S3 method for class 'summary.med_iden'
print(x,...)
```

Arguments

object	a med_iden object created initially call to data.org.
x	a summary.med_iden object created initially call to summary.med_iden
...	other arguments passed to the print function.

Details

summary.med_iden gives a list of identified mediators, covariates and the test results.

Value

The function summary.med_iden return a list of covariates and mediators as identified by a series tests.

mediator	variable names of the identified mediators, either significant in both full model and in relate to the predictor, or being a member of the pre-identified joint mediators.
covariate	variable names of covariates: being significant in the full model but not significantly relate with the predictor.
tests	statistical test results.
results	the original object.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

`"mma"`, `"boot.med.binx"`, `"boot.met.contx"`

Examples

```
data("weight_behavior")
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=FALSE,nu=0.001,
  D=3,distn="bernoulli",family1=binomial(link = "logit"),n2=2)
summary(temp.b.b.glm)
```

summary.mma

Summary of an mma project

Description

Compute the estimations, standard deviations and confidence intervals of the mediation effects.

Usage

```
## S3 method for class 'mma'
summary(object,..., alpha=0.05, plot=TRUE)
## S3 method for class 'summary.mma'
print(x,...)
```

Arguments

object	a mma object created initially call to mma, boot.met.binx, or boot.met.contx.
x	a summary.mma object created initially call to summary.mma.
...	other arguments passed to the print function.
alpha	the alpha level for confidence interval.
plot	default is TRUE, if ture, draw a barplot of the relative mediation effects.

Details

summary.mma gives a list of the estimations and summary statistics based on the bootstrap results. If plot=T, draw a barplot of the relative effects of the direct effect of the predictor and indirect effects of the mediators. Relative effects is defined as the (in)direct effect divided by the total effect. The plot is arranged in order from the largest to the smallest relative effect.

Value

The function `summary.mma` return a list. In each of the following item, `est` is the estimation of the corresponding mediation effects based on the whole data, `mean` is the average estimated effects from the bootstrap samples, `sd` is the standard deviation of the estimates from the bootstrap sample. `upbd` and `lwbd` are the upper and lower bound of the confidence interval of the estimation using parametric method from the bootstrap sample, `upbd_q` and `lwbd_q` are the corresponding quantiles of the estimation from the bootstrap sample.

`ie` a matrix of statistics inference on the indirect effects from the `mma` object.
`te` statistics inference on the total effects from the `mma` object.
`de` statistics inference on the direct effects from the `mma` object.

If `plot=T`, draw a barplot of the relative mediation effects.

Author(s)

Qingzhao Yu <qyu@lsuhsc.edu>

References

Yu, Q., Fan, Y., and Wu, X. (2014). "General Multiple Mediation Analysis With an Application to Explore Racial Disparity in Breast Cancer Survival," *Journal of Biometrics & Biostatistics*,5(2): 189.

See Also

`"mma"`, `"boot.med.binx"`, `"boot.met.contx"`

Examples

```
data("weight_behavior")
x=weight_behavior[,2:14]
y=weight_behavior[,15]
temp.b.b.glm<-mma(x,y,pred=2,contmed=c(8:10,12:13),binmed=c(7,11),
  binref=c(1,1),catmed=6,catref=1,predref="M",alpha=0.4,alpha2=0.4,
  jointm=NULL,margin=1, n=2,seed=sample(1:1000,1),mart=FALSE,nu=0.001,
  D=3,distn="bernoulli",family1=binomial(link = "logit"),n2=2)
summary(temp.b.b.glm)
```

weight_behavior

Weight_Behavior Data Set

Description

This database was obtained from the Louisiana State University Health Sciences Center, New Orleans, by Dr. Richard Scribner. He explored the relationship between BMI and kids behavior through a survey at children, teachers and parents in Grenada in 2014. This data set includes 691 observations and 15 variables.

Usage

```
data(weight_behavior)
```

Format

The data set contains the following variables:

bmi - body mass index, calculated by $\text{weight}(\text{kg})/\text{height}(\text{cm})^2$, numeric

age - children's age in years at the time of survey, numeric

sex - sex of the children, factor

race - race of the children, factor

numpeople - number of people in family, numeric

car - the number of cars in family, numeric

gotosch - the method used to go to school, factor

snack - eat snack or not in a day, binary

tvhours - number of hours watching TV per week, numeric

cmphours - number of hours using computer per week, numeric

cellhours - number of hours playing with cell phones per week, numeric

sports - join in a sport team or not, 1: yes; and 2: no

exercises - number of hours of exercises per week, numeric

sweat - number of hours of sweating activities per week, numeric

overweigh - the child is overweighed or not, binary

Examples

```
data(weight_behavior)
names(weight_behavior)
```


Index

*Topic **Binary Predictor**

boot.med.binx, 3

med.binx, 11

*Topic **Continuous Predictor**

boot.med.contx, 6

med.contx, 13

*Topic **Datasets**

weight_behavior, 23

*Topic **Mediation Analysis**

boot.med.binx, 3

boot.med.contx, 6

med.binx, 11

med.contx, 13

mma, 15

*Topic **Mediator Tests**

data.org, 8

mma, 15

*Topic **Package**

mma-package, 2

boot.med.binx, 2, 3, 7, 12, 15, 18–20, 22, 23

boot.med.contx, 2, 5, 6, 12, 15, 18–20

boot.met.contx, 22, 23

data.org, 2, 8, 10, 18

med.binx, 2, 5, 7, 11, 15, 18

med.contx, 2, 5, 7, 12, 13, 18

mma, 2, 15, 19, 20, 22, 23

mma-package, 2

plot.mma, 18

print.mma, 20

print.summary.med_iden

(summary.med_iden), 21

print.summary.mma (summary.mma), 22

summary.med_iden, 21

summary.mma, 22

weight_behavior, 23