Package: MAZE (via r-universe)

September 8, 2024

Title Mediation Analysis for Zero-Inflated Mediators

Version 0.0.3

Description A novel mediation analysis approach to address zero-inflated mediators containing true zeros and false zeros. See Jiang et al (2023) ``A Novel Causal Mediation Analysis Approach for Zero-Inflated Mediators" <arXiv:2301.10064> for more details.

URL https://github.com/meilinjiang/MAZE

BugReports https://github.com/meilinjiang/MAZE/issues

License GPL-2 | GPL-3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.2

Depends R (>= 4.0), flexmix, numDeriv, pracma

LazyData true

Suggests knitr, rmarkdown

VignetteBuilder knitr

LinkingTo Rcpp

Imports Rcpp, stats, foreach, doParallel, MASS

Repository https://meilinjiang.r-universe.dev

RemoteUrl https://github.com/meilinjiang/maze

RemoteRef HEAD

RemoteSha ae29e73810d9d0a64c516027a70580a31afe8c7e

Contents

DataGen *DataGen*

Description

Generate data under zero-inflated mediation models and calculate the true effects

Usage

```
DataGen(distM, theta, K, num_Z = 0, n, B, x1, x2, zval = NULL, mval = 0)
```

Arguments

distM	distribution of the mediator. One of 'zilonm', 'zinbm', and 'zipm' for zero-inflated log-normal, negative binomial, and Poisson mediators respectively
theta	vector of true parameter values
K	true number of component K in the zero-inflated mixture mediators. Default is $K = 1$ for zero-inflated (non-mixture) mediators
num_Z	number of confounder variables
n	number of observations to generate
B	the upper bound value B to be used in the probability mechanism of observing false zeros
x1	the first value of independent variable of interest
x2	the second value of independent variable of interest
zval	the value of confounders to be conditional on when calculating true effects
mval	the fixed value of mediator to be conditional on when calculating true CDE

Value

true_eff: a vector containing true effects (NIE1, NIE2, NIE, NDE, and CDE)

dat: a data frame containing variables:

- X: independent variable,
- Mobs: observed mediator values (with possibly false zeros)
- M: true mediator values,
- Y: outcome,
- Z: confounder variables (if any)

Author(s)

Meilin Jiang meilin.jiang@ufl.edu and Zhigang Li zhigang.li@ufl.edu

Examples

```
betas.tr <- c(2, 0.12, -6.6, 6.3, -3.8, 0)
delta.tr <- 1.1
alpha0_k.tr <- c(0.4, 1.1)
alpha1_k.tr <- c(0.1, 0.5)
alphas.tr <- rbind(alpha0_k.tr,alpha1_k.tr)
xi0.tr <- -1.5
psi_km1.tr <- c(0.6)
gammas.tr <- c(-1.8, 0.5)
eta.tr <- 1
theta <- c(betas.tr, delta.tr, alphas.tr,
           xi0.tr, psi_km1.tr, gammas.tr, eta.tr)
out <- DataGen(distM = 'zilonm', theta, K = 2, num_Z=0,
               n = 200, B = 20, x1 = 0, x2 = 1, zval = NULL, mval = 0)
(true_eff <- out$true_eff)
dat <- out$dat
```

MAZE *Mediation Analysis for ZEro-inflated mediators*

Description

A novel mediation modeling approach to address zero-inflated mediators containing true zeros and false zeros.

Usage

```
MAZE(
  data,
  distM = c("zilonm", "zinbm", "zipm"),
  K = 1,
  selection = "AIC",
  X,
  M,
  Y,
  Z = NULL,
  XMint = c(TRUE, FALSE),
  x1,
  x2,
  zval = NULL,
  mval = 0,
  B = 20,
  seed = 1,
  ncore = 1
)
```

Arguments

data	a data frame containing variables: an independent variable X, a mediator M, an outcome Y, and confounder variables Z (if any). See example dataset: data(zinb10)
distM	a vector with choices of the distribution of mediator to try with. One or more of 'zilonm', 'zinbm', and 'zipm' for zero-inflated log-normal, negative binomial, and Poisson mediators respectively. Default is c('zilonm', 'zinbm', 'zipm') where all three distributions are fitted and the final mediation model is selected by model selection criterion selection
K	a vector with choices of the number of component K in the zero-inflated mixture mediators to try with. Default is $K = 1$ for zero-inflated (non-mixture) mediators
selection	model selection criterion when more than one model (combination of different values in distM and K) is fitted. Either 'AIC' or 'BIC'. Default is 'AIC'
X	name of the independent variable. Can be continuous or discrete
M	name of the mediator variable. Non-negative values
Y	name of the outcome variable. Continuous values
Z	name(s) of confounder variables (if any)
XMint	a logical vector of length 2 indicating whether to include the two exposure-mediator interaction terms between (i) X and $1_{(M>0)}$ and (ii) X and M. Default is c(TRUE, FALSE), which only includes the first
x1	the first value of independent variable of interest
x2	the second value of independent variable of interest
zval	a vector of value(s) of confounders to be conditional on when estimating effects
mval	the fixed value of mediator to be conditional on when estimating CDE
B	the upper bound value B to be used in the probability mechanism of observing false zeros
seed	an optional seed number to control randomness
ncore	number of cores available for parallel computing

Details

For an independent variable X, a zero-inflated mediator M and a continuous outcome variable Y, the following regression equation is used to model the association between Y and (X, M):

$$Y_{xm1_{(m>0)}} = \beta_0 + \beta_1 m + \beta_2 1_{(m>0)} + \beta_3 x + \beta_4 x 1_{(m>0)} + \beta_5 xm + \epsilon$$

Users can choose to include either one, both, or none of the two exposure-mediator interaction terms between (i) X and $1_{(M>0)}$ and (ii) X and M using the argument XMint.

For mediators, zero-inflated log-normal, zero-inflated negative binomial, and zero-inflated Poisson distributions are considered and can be specified through the argument distM.

The indirect and direct effects (NIE1, NIE2, NIE, NDE, and CDE) are estimated for X changing from x1 to x2. When confounders are present, the conditional effects are estimated given the fixed value zval.

Value

a list containing:

- `results_effects`: a data frame for the results of estimated effects (NIE1, NIE2, NIE, NDE, and CDE). '_cond' for conditional effects at `zval` and '_avg' for average effects
- `results_parameters`: a data frame for the results of model parameters
- `selected_model_name`: a string for the distribution of M and number of components K selected in the final mediation model
- `BIC`: a numeric value for the BIC of the final mediation model
- `AIC`: a numeric value for the AIC of the final mediation model
- `models`: a list with all fitted models
- `analysis2_out`: a list with output from `analysis2()` function (used for internal check)

Author(s)

Meilin Jiang meilin.jiang@ufl.edu and Zhigang Li zhigang.li@ufl.edu

Examples

```
data(zinb10)

maze_out <- MAZE(data = zinb10,
                 distM = c('zilonm', 'zinbm', 'zipm'),  K = 1,
                 selection = 'AIC',
                 X = 'X', M = 'Mobs', Y = 'Y', Z = NULL,
                 XMint = c(TRUE, FALSE),
                 x1 = 0, x2 = 1, zval = NULL, mval = 0,
                 B = 20, seed = 1)
## results of selected mediation model
maze_out$results_effects # indirect and direct effects
maze_out$selected_model_name # selected distribution of the mediator and number of components K
maze_out$results_parameters # model parameters
maze_out$BIC; maze_out$AIC # BIC and AIC of the selected mediation model
```

---

zinb10 *Example dataset 'zinb10'*

Description

An example dataset generated from the proposed model with a zero-inflated negative binomial mediator (K=1). The mediator contains 10% zero values in which half are false zeros.

Usage

```
data(zinb10)
```

Format

An object of class `'data.frame'` with 100 rows and 3 variables:

X independent variable, continuous data type

Y outcome, continuous data type

Mobs observed mediator values with possibly false zeros, count data type

Examples

```
data(zinb10)
head(zinb10)
```

Index