Package ‘BayesGmed’

January 20, 2025

Title Bayesian Causal Mediation Analysis using 'Stan'

Version 0.0.3

Description Performs parametric mediation analysis using the Bayesian g-formula approach for binary and continuous outcomes. The methodology is based on Comment (2018) <doi:10.5281/zenodo.1285275> and a demonstration of its application can be found at Yimer et al. (2022) <doi:10.48550/arXiv.2210.08499>.

Encoding UTF-8

RoxygenNote 7.2.1

Biarch true

Depends R (>= 3.4.0)

Imports methods, Rcpp (>= 0.12.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), rstantools (>= 2.1.1)

LinkingTo BH (>= 1.66.0), Rcpp (>= 0.12.0), RcppEigen (>= 0.3.3.3.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), StanHeaders (>= 2.18.0)

SystemRequirements GNU make

Suggests knitr, rmarkdown

VignetteBuilder knitr

LazyData true

License MIT + file LICENSE

NeedsCompilation yes

Author Belay Birlie Yimer [aut, cre] (<https://orcid.org/0000-0001-8621-6539>),
Mark Lunt [aut],
Marcus Beasley [aut],
Gary J Macfarlane [aut],
John McBeth [aut]

Maintainer Belay Birlie Yimer <belaybirlie.yimer@manchester.ac.uk>

Repository CRAN

Date/Publication 2024-01-31 14:00:06 UTC

Contents

BayesGmed-package *The 'BayesGmed' package.*

Description

A DESCRIPTION OF THE PACKAGE

References

Stan Development Team (2022). RStan: the R interface to Stan. R package version 2.21.5. https://mc-stan.org

bayesgmed *Estimate a causal mediation effects*

Description

Estimates various quantities for causal mediation analysis using 'Stan'.

Usage

```
bayesgmed(
  outcome,
  mediator,
  treat,
  covariates = NULL,
  dist.y = "continuous",
  dist.m = "continuous",
  link.y = "identity",
  link.m = "identity",
  data,
  priors = NULL,
  ...
)
```

Arguments

`outcome` a character string indicating the name of the outcome variable.

`mediator` a character string indicating the name of the mediator variable.

`treat` a character string indicating the name of the treatment variable. The treatment variable is considered binary and should be coded as 0 for control and 1 for treated.

`covariates` a character vector indicating the name of the confounding variables.

`dist.y` a character string indicating the family distribution of the outcome. E.g., dist.y = "bernoulli" will fit a logistic regression for the outcome.

`dist.m` a character string indicating the family distribution of the mediator, E.g., dist.m = "bernoulli" will fit a logistic regression model for the mediator

`link.y` a character string indicating the link function to be used for the outcome model.

`link.m` a character string indicating the link function to be used for the mediator model.

`data` A `data.frame` or a `tibble` object.

`priors` A list of named values for the prior scale parameters. See details.

`...` Other optional parameters passed to `rstan::stan()`.

Details

This is the main function for estimating causal mediation effects from several types of data within the Bayesian framework. We followed the potential outcome framework for effects definition and the package uses the 'rstan' utility functions for exploring the posterior distribution.

priors:

Users may pass a list of named values for the priors argument. The values will be used to define the scale parameter of the respective prior distributions. This list may specify some or all of the following parameters: priors <- list(scale_m = 2.5*diag(P_m) scale_y = 2.5*diag(P_y), location_m = rep(0, P_m) location_y = rep(0, P_y), scale_sd_y = 2.5, scale_sd_m = 2.5) where P_m is the number of regression parameters (including the intercept) in the mediator model and P_y is the number of regression parameters in the outcome model.

Value

An object of 'S4' class 'stanfit', with all its available methods.

Author(s)

Belay Birlie Yimer <`belaybirlie.yimer@manchester.ac.uk`>

References

1. McCandless, L.C. and J.M. Somers, *Bayesian sensitivity analysis for unmeasured confounding in causal mediation analysis.* Statistical Methods in Medical Research, 2019. (28)(2): p. 515-531.
2. Comment, L., Coull, B. A., Zigler, C., and Valeri, L. (2019). Bayesian data fusion for unmeasured confounding. arXiv preprint arXiv:1902.10613.

Examples

```
## Run example using the example_data
data(example_data)
fit1 <- bayesgmed(outcome = "Y", mediator =  "M", treat = "A", covariates = c("Z1", "Z2"),
 dist.y = "binary", dist.m = "binary", link.y = "logit", link.m = "logit", data = example_data)
bayesgmed_summary(fit1)

# With priors
P <- 3 # number of covariates plus the intercept term
priors <- list(scale_m = 2.5*diag(P+1), scale_y = 2.5*diag(P+2),
location_m = rep(0, P+1), location_y = rep(0, P+2))
fit1 <- bayesgmed(outcome = "Y", mediator =  "M", treat = "A", covariates = c("Z1", "Z2"),
dist.y = "binary", dist.m = "binary", link.y = "logit", link.m = "logit", priors = priors,
data = example_data)
bayesgmed_summary(fit1)
```

bayesgmed_sens *Conduct sensitivity analysis for unmeasured confounder*

Description

'bayesgmed_sens' is used to conduct sensitivity analysis for unmeasured confounders in mediation analysis.

Usage

```
bayesgmed_sens(
  outcome,
  mediator,
  treat,
  covariates = NULL,
  dist.y = "continuous",
  dist.m = "continuous",
  link.y = "identity",
  link.m = "identity",
  data,
  priors = NULL,
  ...
)
```

Arguments

`outcome` a character string indicating the name of the outcome variable.

`mediator` a character string indicating the name of the mediator variable.

`treat` a character string indicating the name of the treatment variable. The treatment variable is considered binary and should be coded as 0 for control and 1 for treated.

`covariates` a character vector indicating the name of the confounding variables.

`dist.y` a character string indicating the family distribution of the outcome. E.g., dist.y = "bernoulli" will fit a logistic regression the outcome.

`dist.m` a character string indicating the family distribution of the mediator E.g., dist.m = "bernoulli" will fit a logistic regression the mediator

`link.y` a character string indicating the link function to be used for the outcome model.

`link.m` a character string indicating the link function to be used for the mediator model.

`data` A `data.frame` or a `tibble` object.

`priors` A list of named values for the prior scale parameters. See details.

`...` Other optional parameters passed to `rstan::stan()`.

Details

Perform a sensitivity analysis for unmeasured confounding following the approach proposed by McCandless LC and Somers JM (2019). This is done by incorporating uncertainty about unmeasured confounding in the outcome and mediator model through a prior distribution. One can control the size of unmeasured confounding by varying the location and scale parameters of the prior distribution values for the bias parameter (i.e., gamma). See the Vignette for more details.

priors:

Users may pass a list of named values for the priors argument. The values will be used to define the scale parameter of the respective prior distributions. This list may specify some or all of the following parameters: priors <- list(scale_m = 2.5*diag(P_m) scale_y = 2.5*diag(P_y), llocation_m = rep(0, P_m) location_y = rep(0, P_y), location_gamma = rep(0, 4), scale_gamma = 0.1*diag(4), scale_sd_y = 2.5, scale_sd_m = 2.5) where `P_m` is the number of regression parameters (including the intercept) in the mediator model and `P_y` is the number of regression parameters in the outcome model. Note that there are 4 bias parameters i.e., gamma).

Value

An object of 'S4' class 'stanfit', with all its available methods.

Author(s)

Belay Birlie Yimer `<belaybirlie.yimer@manchester.ac.uk>`

References

1. McCandless, L.C. and J.M. Somers, *Bayesian sensitivity analysis for unmeasured confounding in causal mediation analysis.* Statistical Methods in Medical Research, 2019. (28)(2): p. 515-531.
2. Comment, L., Coull, B. A., Zigler, C., and Valeri, L. (2019). Bayesian data fusion for unmeasured confounding. arXiv preprint arXiv:1902.10613.

Examples

```
## Run example using the example_data
data(example_data)
# priors including for unmeasured single confounder
P <- 3 # number of covariates plus the intercept term
priors <- list(scale_m = 2.5*diag(P+1), scale_y = 2.5*diag(P+2), location_m = rep(0, P+1),
location_y = rep(0, P+2), location_gamma = rep(0,4), scale_gamma = 0.5*diag(4))
fit1 <- bayesgmed_sens(outcome = "Y", mediator =  "M", treat = "A", covariates = c("Z1", "Z2"),
dist.y = "binary", dist.m = "binary", link.y = "logit", link.m = "logit", priors = priors,
data = example_data)
bayesgmed_summary(fit1)
```

---

bayesgmed_summary *Print a summary of the estimated causal mediation model*

Description

Print a summary of the estimated causal mediation model

Usage

```
bayesgmed_summary(
  model = NULL,
  level = 0.95,
 pars = c("NDE_control", "NDE_treated", "NIE_control", "NIE_treated", "ANDE", "ANIE",
    "TE"),
  digits = 3
)
```

Arguments

model	A stanfit object obtained from bayesgmed().
level	The "confidence" level that defines the limits of the credible intervals (default is .95, i.e. 95% CIs).
pars	The parameters to summarize (defaults to main causal effect estimands similar to the mediation package; see Details).
digits	The number of decimal points to display in the output (default is 3).

Details

After estimating a model with bayesgmed(), use bayesgmed_summary(fit) to display the estimated results, where fit is an object containing the fitted model. By default, bayesgmed_summary() only displays a subset of the estimated parameters: - NDE_control: direct effect estimate when the exposure level is set to the control value. - NDE_treated: direct effect estimate when the exposure level is set to the treated value. - NIE_control: mediated effect estimate when the exposure level is set to the control value. - NIE_treated: mediated effect estimate when the exposure level is set to the treated value. - ANDE: average direct effect of X on Y. - ANIE: average indirect effect of X on Y. - TE: the total effect of A on Y. To display all estimated parameters where all chains merged, set pars = NULL. This will print all parameters defined in the model definitions, including the most important ones: - alphaZ[]: parameter estimate of the confounders (i.e., X -> Y) relationship, listed in the order they are specified in the covariates argument of bayesgmed(). alpha[1] is the intercept. - alphaM: parameter estimate of the M -> Y relationship. - alphaA: parameter estimate of the A -> Y relationship. - betaZ: parameter estimate of the confounders (i.e., X -> M) relationship, listed in the order they are specified in the covariates argument of bayesgmed(). beta[1] is the intercept

To learn more about the additional parameters, refer to the Stan code (cat(get_stancode(fit))).

Value

A data.frame summarizing the estimated causal mediation model, including the following columns:

- Parameter: the name of the parameter.
- Mean: the mean of the parameter's posterior distribution.
- Median: the median of the parameter's posterior distribution.
- SE: the standard deviation of the parameter's posterior distribution.
- ci_lwr: the lower limit of the credible interval.
- ci_upr: the upper limit of the credible interval.
- n_eff: the number of efficient samples.
- Rhat: a value of 1.00 suggests model convergence.

Author(s)

Belay B. Yimer <belaybirlie.yimer@manchester.ac.uk>

example_data *example_data: A simulated data set for illustrating mediation analysis in BayesGmed.*

Description

The data set contains a binary outcome variable Y, a binary exposure A, a binary mediator M, and two confounding variables (Z1 and Z2) for 300 individuals.

Usage

```
example_data
```

Format

`example_data`:

A data frame with 300 rows and 5 columns:

Z1, Z2 confounders

A binary exposure

M binary mediator

Y binary outcome ...

Index