Package 'mrbayes'

Type Package

Title Bayesian Summary Data Models for Mendelian Randomization Studies

Version 0.2.0

Description Bayesian estimation of inverse variance weighted (IVW), Burgess et al. (2013) <doi:10.1002/gepi.21758>, and MR-Egger, Bowden et al. (2015) <doi:10.1093/ije/dyv080>, summary data models for Mendelian randomization analyses.

Depends R (>= 3.5.0)

Imports methods, Rcpp (>= 0.12.0), rstan (>= 2.18.1), rstantools (>= 2.0.0)

License GPL-3

Encoding UTF-8

LazyData true

Suggests testthat, rjags (>= 4-9), tidyverse, knitr, ggplot2, ggmcmc, foreign, DescTools, bayesmeta, utils, kableExtra, roxygen2, bookdown

URL https://github.com/okezie94/mrbayes

BugReports https://github.com/okezie94/mrbayes/issues

RoxygenNote 7.1.0

SystemRequirements GNU make

Biarch true

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

NeedsCompilation yes

Author Okezie Uche-Ikonne [aut, cre] (<https://orcid.org/0000-0002-8461-8099>),
Frank Dondelinger [aut] (<https://orcid.org/0000-0003-1816-6300>),
Tom Palmer [aut] (<https://orcid.org/0000-0003-4655-4511>)

Maintainer Okezie Uche-Ikonne <o.uche-ikonne@lancaster.ac.uk>

Repository CRAN

R topics documented:

---

mrbayes-package   *mrbayes: Bayesian implementation of the IVW and MR-Egger models for two-sample Mendelian randomization analyses*

Description

Bayesian implementation of the IVW and MR-Egger models for two-sample Mendelian randomization analyses.

References

Stan Development Team (2019). RStan: the R interface to Stan. R package version 2.19.2. https://mc-stan.org

---

bmi_insulin   *Dataset from Richmond et. al 2017 investigating the association of BMI on insulin resistance*

Description

A two-sample summary level dataset, Richmond et al. (2017) https://dx.doi.org/10.1101/155739, containing 14 single nucleiodtide polymorphisms (SNPs) which have genotype-phenotype associations (BMI) and genotype-outcome associations (insulin) with their respective standard errors.

Usage

```
bmi_insulin
```

Format

A data frame with 14 rows and 44 columns:

rsid SNP RSID number

beta.exposure The genotype-BMI associations

beta.outcome The genotype-outcome associations

se.exposure The standard errors of the genotype-phenotype associations

se.outcome The standard errors of the genotype-outcome associations

Details

bmi_insulin.

References

Richmond, R. et al., Investigating the role of insulin in increased adiposity: Bi-directional Mendelian randomization study. bioRxiv, 2017, https://dx.doi.org/10.1101/155739.

---

do_data	*Dataset from Do et al., Nat Gen, 2013 containing summary level data on associations of genotypes with lipid traits and the risk of coronary heart diseases*

Description

A summary-level dataset, from Do et al. (2013) https://dx.doi.org/10.1038/ng.2795, containing 185 single nucleodtide polymorphisms (SNPs) which have genotype-phenotype associations and standard errors for low-density lipoprotein cholestrol, and genotype-outcome associations for coronary heart disease with their respective standard errors.

Usage

```
do_data
```

Format

A data frame with 185 rows and 9 columns:

rsid SNPs RSID number

beta.exposure The genotype-phenotype associations for low-density lipoprotein cholestrol

beta.outcome The genotype-outcome associations in this case the outcome is coronary heart disease

se.exposure The standard errors of the genotype-low-density lipoprotein cholestrol associations

se.outcome The standard errors of the genotype-outcome coronary heart disease associations

Details

do_data.

References

Do, R. et al., Common variants associated with plasma triglycerides and risk for coronary artery disease. Nature Genetics, 2013, 45, 1345-1352, https://dx.doi.org/10.1038/ng.2795.

---

mr_egger_rjags	*Bayesian implementation of the MR-Egger model with choice of prior distributions fitted using JAGS.*

Description

Bayesian implementation of the MR-Egger model with choice of prior distributions fitted using JAGS.

Usage

```
mr_egger_rjags(
  object,
  prior = "default",
  betaprior = "",
  sigmaprior = "",
  n.chains = 3,
  n.burn = 1000,
  n.iter = 5000,
  seed = NULL,
  rho = 0.5,
  ...
)
```

Arguments

object	A data object of class mr_format.
prior	A character string for selecting the prior distributions;

- "default" selects a non-informative set of priors;
- "weak" selects weakly informative priors;
- "pseudo" selects a pseudo-horseshoe prior on the causal effect;
- "joint" selects a joint prior on the intercept and slope.

betaprior	A character string in JAGS syntax to allow a user defined prior for the causal effect.
sigmaprior	A character string in JAGS syntax to allow a user defined prior for the residual standard deviation.

n.chains	Numeric indicating the number of chains used in the MCMC estimation, the default is 3 chains.
n.burn	Numeric indicating the burn-in period of the Bayesian MCMC estimation. The default is 1000 samples.
n.iter	Numeric indicating the number of iterations in the Bayesian MCMC estimation. The default is 5000 iterations.
seed	Numeric indicating the random number seed. The default is the rjags default.
rho	Numeric indicating the correlation coefficient input into the joint prior distribution. The default value is 0.5.
...	Additional arguments passed through to rjags::jags.model().

Value

An object of class eggerjags containing the following components:

AvgPleio The mean of the simulated pleiotropic effect

CausalEffect The mean of the simulated causal effect

StandardError Standard deviation of the simulated causal effect

sigma The value of the residual standard deviation

CredibleInterval The credible interval for the causal effect, which includes the lower (2.5%), median (50%) and upper intervals (97.5%)

samples Output of the Bayesian MCMC samples

Priors The specified priors

References

Bowden et. al., Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International Journal of Epidemiology 2015. 44(2): p. 512-525. https://doi.org/10.1093/ije/dyv080

Examples

```
fit <- mr_egger_rjags(bmi_insulin)
summary(fit)
plot(fit$samples)
# 90% credible interval
fitdf <- do.call(rbind.data.frame, fit$samples)
cri90 <- quantile(fitdf$Estimate, probs = c(0.05,0.95))
print(cri90)
```

---

mr_egger_stan   *Bayesian inverse variance weighted model with a choice of prior distributions fitted using Stan*

Description

Bayesian inverse variance weighted model with a choice of prior distributions fitted using Stan.

Usage

```
mr_egger_stan(
  data,
  prior = 1,
  n.chains = 3,
  n.burn = 1000,
  n.iter = 5000,
  seed = 12345,
  rho = 0.5,
  ...
)
```

Arguments

data	A data of class [mr_format]().
prior	An integer for selecting the prior distributions;
	• 1 selects a non-informative set of priors;
	• 2 selects weakly informative priors;
	• 3 selects a pseudo-horseshoe prior on the causal effect;
	• 4 selects joint prior of the intercept and causal effect estimate.
n.chains	Numeric indicating the number of chains used in the HMC estimation in rstan, the default is 3 chains.
n.burn	Numeric indicating the burn-in period of the Bayesian HMC estimation. The default is 1000 samples.
n.iter	Numeric indicating the number of iterations in the Bayesian HMC estimation. The default is 5000 iterations.
seed	Numeric indicating the random number seed. The default is 12345.
rho	Numeric indicating the correlation coefficient input into the joint prior distribution. The default is 0.5.
...	Additional arguments passed through to [rstan::sampling()]().

Value

An object of class [stanfit]().

References

Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International Journal of Epidemiology, 2015, 44, 2, 512-525. <http://dx.doi.org/10.1093/ije/dyv080>.

Stan Development Team (2020). "RStan: the R interface to Stan." R package version 2.19.3, <http://mc-stan.org/>.

Examples

```
# Note we recommend setting n.burn and n.iter to larger values
egger_fit <- mr_egger_stan(bmi_insulin, n.burn = 500, n.iter = 1000)
print(egger_fit)
```

mr_format	*Organises the summary level data for use in the Bayesian MR functions*

Description

Organises the summary level data for use in the Bayesian MR functions

Usage

```
mr_format(rsid, xbeta, ybeta, xse, yse)
```

Arguments

rsid	A vector of genetic variants used for analysis, if unspecified a vector is automatically generated.
xbeta	A numeric vector of the instrument-phenotype associations.
ybeta	A numeric vector of the instrument-outcome associations.
xse	The standard errors of the instrument-phenotype associations xbeta.
yse	The standard errors of the instrument-outcome associations ybeta.

Value

A formatted data frame for analysis of class mr_format.

Examples

```
data(bmi_insulin)
dat <- mr_format(rsid = bmi_insulin[,"rsid"],
          xbeta = bmi_insulin[,"beta.exposure"],
          ybeta = bmi_insulin[,"beta.outcome"],
          xse = bmi_insulin[,"se.exposure"],
          yse = bmi_insulin[,"se.outcome"])
class(dat)
```

---

mr_ivw_rjags   *Bayesian inverse variance weighted model with a choice of prior dis-*
                                tributions fitted using JAGS.

Description

Bayesian inverse variance weighted model with a choice of prior distributions fitted using JAGS.

Usage

```
mr_ivw_rjags(
  object,
  prior = "default",
  betaprior = "",
  n.chains = 3,
  n.burn = 1000,
  n.iter = 5000,
  seed = NULL,
  ...
)
```

Arguments

object	A data object of class `mr_format`.
prior	A character string for selecting the prior distributions;
	• `"default"` selects a non-informative set of priors;
	• `"weak"` selects weakly informative priors;
	• `"pseudo"` selects a pseudo-horseshoe prior on the causal effect.
betaprior	A character string in JAGS syntax to allow a user defined prior for the causal effect.
n.chains	Numeric indicating the number of chains used in the MCMC estimation, the default is 3 chains.
n.burn	Numeric indicating the burn-in period of the Bayesian MCMC estimation. The default is 1000 samples.
n.iter	Numeric indicating the number of iterations in the Bayesian MCMC estimation. The default is 5000 iterations.

seed	Numeric indicating the random number seed. The default is the rjags default.
...	Additional arguments passed through to rjags::jags.model().

Value

An object of class ivwjags containing the following components:

CausalEffect The mean of the simulated causal effects

StandardError Standard deviation of the simulated causal effects

CredibleInterval The credible interval for the causal effect, which indicates the lower(2.5%), median (50%) and upper intervals (97.5%)

samples Output of the Bayesian MCMC samples with the different chains

Priors The specified priors

References

Burgess, S., Butterworth, A., Thompson S.G. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic Epidemiology, 2013, 37, 7, 658-665 https://dx.doi.org/10.1002/gepi.21758.

Examples

```
fit <- mr_ivw_rjags(bmi_insulin)
print(fit)
summary(fit)
plot(fit$samples)
# 90% credible interval
fitdf <- do.call(rbind.data.frame, fit$samples)
cri90 <- quantile(fitdf$Estimate, probs = c(0.05,0.95))
print(cri90)
```

---

mr_ivw_stan	*Bayesian inverse variance weighted model with a choice of prior distributions fitted using RStan.*

Description

Bayesian inverse variance weighted model with a choice of prior distributions fitted using RStan.

Usage

```
mr_ivw_stan(
  data,
  prior = 1,
  n.chains = 3,
  n.burn = 1000,
```

```
  n.iter = 5000,
  seed = 12345,
  ...
)
```

Arguments

`data`	A data of class `mr_format`.
`prior`	An integer for selecting the prior distributions;

- 1 selects a non-informative set of priors;
- 2 selects weakly informative priors;
- 3 selects a pseudo-horseshoe prior on the causal effect.

`n.chains`	Numeric indicating the number of chains used in the HMC estimation in rstan, the default is 3 chains.
`n.burn`	Numeric indicating the burn-in period of the Bayesian HMC estimation. The default is `1000` samples.
`n.iter`	Numeric indicating the number of iterations in the Bayesian MCMC estimation. The default is `5000` iterations.
`seed`	Numeric indicating the random number seed. The default is `12345`.
`...`	Additional arguments passed through to `rstan::sampling()`.

Value

An object of class `stanfit`.

References

Burgess, S., Butterworth, A., Thompson S.G. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic Epidemiology, 2013, 37, 7, 658-665 https://dx.doi.org/10.1002/gepi.21758.

Stan Development Team (2020). "RStan: the R interface to Stan." R package version 2.19.3, http://mc-stan.org/.

Examples

```
ivw_fit <- mr_ivw_stan(bmi_insulin)
print(ivw_fit)
rstan::traceplot(ivw_fit)
```

---

mr_radialegger_rjags   *Bayesian radial MR-Egger model with a choice of prior distributions fitted using JAGS.*

Description

Bayesian radial MR-Egger model with a choice of prior distributions fitted using JAGS.

Usage

```
mr_radialegger_rjags(
  object,
  prior = "default",
  betaprior = "",
  sigmaprior = "",
  n.chains = 3,
  n.burn = 1000,
  n.iter = 5000,
  seed = NULL,
  rho = 0.5,
  ...
)
```

Arguments

object	A data object of class `mr_format`.
prior	A character string for selecting the prior distributions;
	• `"default"` selects a non-informative set of priors;
	• `"weak"` selects weakly informative priors;
	• `"pseudo"` selects a pseudo-horseshoe prior on the causal effect;
	• `"joint"` selects a joint prior on the intercept and slope.
betaprior	A character string in JAGS syntax to allow a user defined prior for the causal effect.
sigmaprior	A character string in JAGS syntax to allow a user defined prior for the residual standard deviation.
n.chains	Numeric indicating the number of chains used in the MCMC estimation, the default is 3 chains.
n.burn	Numeric indicating the burn-in period of the Bayesian MCMC estimation. The default is `1000` samples.
n.iter	Numeric indicating the number of iterations in the Bayesian MCMC estimation. The default is `5000` iterations.
seed	Numeric indicating the random number seed. The default is the rjags default.
rho	Numeric indicating the correlation coefficient input into the joint prior distribution. The default is `0.5`.
...	Additional arguments passed through to `rjags::jags.model()`.

Value

An object of class `radialeggerjags` containing the following components:

AvgPleio The mean of the simulated pleiotropic effect

CausalEffect The mean of the simulated causal effect

StandardError Standard deviation of the simulated causal effect

sigma The mean of the simaulted residual standard deviation

CredibleInterval The credible interval for the causal effect, which includes the lower (2.5%), median (50%) and upper intervals (97.5%)

samples Output of the Bayesian MCMC samples

Prior The specified priors

References

Bowden, J., et al., Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. International Journal of Epidemiology, 2018. 47(4): p. 1264-1278. https://doi.org/10.1093/ije/dyy101.

Examples

```
fit <- mr_radialegger_rjags(bmi_insulin)
summary(fit)
plot(fit$samples)
# 90% credible interval
fitdf <- do.call(rbind.data.frame, fit$samples)
cri90 <- quantile(fitdf$Estimate, probs = c(0.05,0.95))
print(cri90)
```

---

mr_radialegger_stan   *Bayesian inverse variance weighted model with a choice of prior distributions fitted using RStan.*

Description

Bayesian inverse variance weighted model with a choice of prior distributions fitted using RStan

Usage

```
mr_radialegger_stan(
  data,
  prior = 1,
  n.chains = 3,
  n.burn = 1000,
  n.iter = 5000,
  rho = 0.5,
```

```
    seed = 12345,
    ...
)
```

Arguments

data	A data of class `mr_format`.
prior	An integer for selecting the prior distributions;

  - 1 selects a non-informative set of priors;
  - 2 selects weakly informative priors;
  - 3 selects a pseudo-horseshoe prior on the causal effect;
  - 4 selects joint prior of the intercept and causal effect estimate.

n.chains	Numeric indicating the number of chains used in the HMC estimation in rstan, the default is 3 chains.
n.burn	Numeric indicating the burn-in period of the Bayesian HMC estimation. The default is 1000 samples.
n.iter	Numeric indicating the number of iterations in the Bayesian HMC estimation. The default is 5000 iterations.
rho	Numeric indicating the correlation coefficient input into the joint prior distribution. The default is 0.5.
seed	Numeric indicating the random number seed. The default is 12345.
...	Additional arguments passed through to `rstan::sampling()`.

Value

An object of class `stanfit`.

References

Bowden, J., et al., Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. International Journal of Epidemiology, 2018. 47(4): p. 1264-1278. https://doi.org/10.1093/ije/dyy101.

Stan Development Team (2020). "RStan: the R interface to Stan." R package version 2.19.3, http://mc-stan.org/.

Examples

```
# Note we recommend setting n.burn and n.iter to larger values
radegger_fit <- mr_radialegger_stan(bmi_insulin, n.burn = 500, n.iter = 1000)
print(radegger_fit)
```

Index