Package 'MRcML'

August 23, 2024

Title Mendelian randomization method with constrained maximum likelihood.

Version 0.0.0.9000

Description Perform MRcML method.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

Author Haoran Xue [aut, cre],
Xiaotong Shen [aut],
Wei Pan [aut]

Maintainer Haoran Xue <xuexx268@umn.edu>

Depends R (>= 2.10)

R topics documented:

---

cML_estimate   *Estimate with Regular Likelihood*

Description

Estimate theta, b vector, r vector with constrained maximum likelihood.

1

Usage

```
cML_estimate(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K,
  initial_theta = 0,
  initial_mu = rep(0, length(b_exp)),
  maxit = 100
)
```

Arguments

b_exp	Vector of estimated effects for exposure.
b_out	Vector or estimated effects for outcome.
se_exp	Vector of standard errors for exposure.
se_out	Vector of standard errors for outcome.
K	Constraint parameter, number of invalid IVs.
initial_theta	Starting point for theta.
initial_mu	Starting point for mu.
maxit	Maximum number of iteration.

Value

A list contains: theta is the estimate causal effect, b_vec is the estimated vector of b, r_vec is the estimated vector of r.

---

cML_estimate_DP	*Estimate with With Data Perturbation*

Description

With multiple perturbed data, get estimated theta, se of estimated theta and negative log-likelihood, using multiple random starting points.

Usage

```
cML_estimate_DP(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K,
  num_pert = 200,
  random_start = 0,
  maxit = 100
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K`	Constraint parameter, number of invalid IVs.
`num_pert`	Number of perturbation, default is 200.
`random_start`	Number of random starting points, default is 0.
`maxit`	Maximum number of iteration.

Value

A list contains: theta_v is the vector estimated thetas, se_v is vector of standard errors of estimated thetas, l_v is vector of negative log-likelihood. Vectors all have length num_pert.

---

cML_estimate_random   *Estimate with Regular Likelihood Using Multiple Random Start Points*

Description

Get estimated theta, se of estimated theta and negative log-likelihood, using multiple random starting points.

Usage

```
cML_estimate_random(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K,
  random_start = 0,
  maxit = 100
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K`	Constraint parameter, number of invalid IVs.
`random_start`	Number of random starting points, default is 0.
`maxit`	Maximum number of iteration.

Value

A list contains: theta is the estimate causal effect, se is standard error of estimated theta, l is negative log-likelihood, r_est is estimated r vector.

cML_SdTheta   *Standard Error of Estimated Theta*

Description

Get the standard error of estimated theta from constrained maximum likelihood.

Usage

```
cML_SdTheta(b_exp, b_out, se_exp, se_out, theta, b_vec, r_vec)
```

Arguments

b_exp	Vector of estimated effects for exposure.
b_out	Vector or estimated effects for outcome.
se_exp	Vector of standard errors for exposure.
se_out	Vector of standard errors for outcome.
theta	Estimated theta from cML.
b_vec	Estimated vector of b from cML.
r_vec	Estimated vector of r from cML.

Value

Standard error of theta.

mr_cML   *MRcML method*

Description

This is the main function of MRcML method, without data perturbation.

Usage

```
mr_cML(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K_vec = 0:(length(b_exp) - 2),
  random_start = 0,
  maxit = 100,
  random_seed = 0,
  n
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K_vec`	Sets of candidate K's, the constraint parameter representing number of invalid IVs.
`random_start`	Number of random start points for cML, default is 0.
`maxit`	Maximum number of iterations for each optimization.
`random_seed`	Random seed, an integer. Default is 0, which does not set random seed; user could specify a positive integer as random seed to get replicable results.
`n`	Sample size.

Value

A list contains full results of cML methods. MA_BIC_theta, MA_BIC_se, MA_BIC_p: Estimate of theta, its standard error and p-value from cML-MA-BIC. Similarly for BIC_theta, BIC_se, BIC_p from cML-BIC; for MA_AIC_theta, MA_AIC_se, MA_AIC_p from cML-MA-AIC; for AIC_DP_theta, AIC_DP_se, AIC_DP_p from cML-AIC. BIC_invalid is the set of invalid IVs selected by cML-BIC, AIC_invalid is the set of invalid IVs selected by cML-AIC. BIC_vec is the BIC vector.

`mr_cML_DP`	*MRcML method with Data Perturbation*

Description

This is the main function of MRcML method with data perturbation.

Usage

```
mr_cML_DP(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K_vec = 0:(length(b_exp) - 2),
  random_start = 0,
  random_start_pert = 0,
  maxit = 100,
  num_pert = 200,
  random_seed = 0,
  n
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K_vec`	Sets of candidate K's, the constraint parameter representing number of invalid IVs.
`random_start`	Number of random start points for cML, default is 0.
`random_start_pert`	
	Number of random start points for cML with data perturbation, default is 0.
`maxit`	Maximum number of iterations for each optimization.
`num_pert`	Number of perturbation, default is 200.
`random_seed`	Random seed, an integer. Default is 0, which does not set random seed; user could specify a positive integer as random seed to get replicable results.
`n`	Sample size.

Value

A list contains full results of cML methods. MA_BIC_theta, MA_BIC_se, MA_BIC_p: Estimate of theta, its standard error and p-value from cML-MA-BIC. Similarly for BIC_theta, BIC_se, BIC_p from cML-BIC; for MA_BIC_DP_theta, MA_BIC_DP_se, MA_BIC_DP_p from cML-MA-BIC-DP; for BIC_DP_theta, BIC_DP_se, BIC_DP_p from cML-BIC-DP. BIC_invalid is the set of invalid IVs selected by cML-BIC.

---

mr_cML_DP_Overlap   *MRcML method for overlapping samples with Data Perturbation*

Description

This is the main function of MRcML method for overlapping samples with data perturbation.

Usage

```
mr_cML_DP_Overlap(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K_vec = 0:(length(b_exp) - 2),
  random_start = 0,
  random_start_pert = 0,
  maxit = 100,
  num_pert = 100,
  random_seed = 0,
  n,
  rho = 0
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K_vec`	Sets of candidate K's, the constraint parameter representing number of invalid IVs.
`random_start`	Number of random start points for cML, default is 0.
`random_start_pert`	
	Number of random start points for cML with data perturbation, default is 0.
`maxit`	Maximum number of iterations for each optimization.
`num_pert`	Number of perturbation, default is 200.
`random_seed`	Random seed, an integer. Default is 0, which does not set random seed; user could specify a positive integer as random seed to get replicable results.
`n`	Sample size.
`rho`	Correlation between GWAS summary statistics due to overlapping samples.

Value

A list contains full results of cML methods. MA_BIC_theta, MA_BIC_se, MA_BIC_p: Estimate of theta, its standard error and p-value from cML-MA-BIC. Similarly for BIC_theta, BIC_se, BIC_p from cML-BIC; for MA_BIC_DP_theta, MA_BIC_DP_se, MA_BIC_DP_p from cML-MA-BIC-DP; for BIC_DP_theta, BIC_DP_se, BIC_DP_p from cML-BIC-DP. BIC_invalid is the set of invalid IVs selected by cML-BIC.

`mr_cML_Overlap`	*MRcML method for overlapping samples*

Description

This is the main function of MRcML method with overlapping samples, without data perturbation.

Usage

```
mr_cML_Overlap(
  b_exp,
  b_out,
  se_exp,
  se_out,
  K_vec = 0:(length(b_exp) - 2),
  random_start = 0,
  maxit = 100,
  random_seed = 0,
  n,
  rho = 0
)
```

Arguments

`b_exp`	Vector of estimated effects for exposure.
`b_out`	Vector or estimated effects for outcome.
`se_exp`	Vector of standard errors for exposure.
`se_out`	Vector of standard errors for outcome.
`K_vec`	Sets of candidate K's, the constraint parameter representing number of invalid IVs.
`random_start`	Number of random start points for cML, default is 0.
`maxit`	Maximum number of iterations for each optimization.
`random_seed`	Random seed, an integer. Default is 0, which does not set random seed; user could specify a positive integer as random seed to get replicable results.
`n`	Sample size.
`rho`	Correlation between GWAS summary statistics due to overlapping samples.

Value

A list contains full results of cML methods. MA_BIC_theta, MA_BIC_se, MA_BIC_p: Estimate of theta, its standard error and p-value from cML-MA-BIC. Similarly for BIC_theta, BIC_se, BIC_p from cML-BIC. BIC_invalid is the set of invalid IVs selected by cML-BIC, BIC_vec is the BIC vector.

Index