

Package ‘RsqMed’

January 20, 2025

Title Total Mediation Effect Size Measure for High-Dimensional Mediators

Version 1.1

Description An implementation of calculating the R-squared measure as a total mediation effect size measure and its confidence interval for moderate- or high-dimensional mediator models. It gives an option to filter out non-mediators using variable selection methods. The original R package is directly related to the paper Yang et al (2021) ``Estimation of mediation effect for high-dimensional omics mediators with application to the Framingham Heart Study" <doi:10.1101/774877>. The new version contains a choice of using cross-fitting, which is computationally faster. The details of the cross-fitting method are available in the paper Xu et al (2023) ``Speeding up interval estimation for R2-based mediation effect of high-dimensional mediators via cross-fitting" <doi:10.1101/2023.02.06.527391>.

License GPL-3

Depends R (>= 3.5.0)

Imports SIS (>= 0.8), GMMAT (>= 1.4.1)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

NeedsCompilation no

Maintainer Tianzhong Yang <yang3704@umn.edu>

Repository CRAN

Author Tianzhong Yang [aut, cre],
Chunlin Li [aut],
Zhichao Xu [ctb]

Date/Publication 2023-11-14 17:20:12 UTC

Contents

CF_Rsq.measure	2
CI.Rsq.measure	3
example	5
Rsq.measure	5

CF_Rsq.measure	<i>Function to calculate the Rsq function as a total effect size measure for mediation effect using cross-fitted estimation</i>
----------------	---

Description

Function to calculate the Rsq function as a total effect size measure for mediation effect using cross-fitted estimation

Usage

```
CF_Rsq.measure(
  Y,
  M,
  Covar = NULL,
  X,
  iter.max = 3,
  nsis = NULL,
  first.half = TRUE,
  seed = 2022,
  tune = c("aic", "bic"),
  penalty = c("MCP", "lasso")
)
```

Arguments

Y	vector of the outcome of interest; outcome has to follow a Gaussian distribution.
M	matrix of putative mediators
Covar	covariates matrix
X	vector of the independent variable of interest, e.g. environmental variable
iter.max	Maximum number of iterations used in iSIS, default = 3 (details see the SIS package).
nsis	Number of predictors recruited by iSIS, default = NULL
first.half	TRUE: split sample into two halves by the order in the dataset. FALSE: randomly split samples into halves, default = TRUE.
seed	Random seed used for sample splitting, default = 2022.
tune	Method for tuning the regularization parameter of the penalized likelihood subproblems and of the final model selected by (i)SIS. Options include tune = 'bic' and tune = 'aic'.
penalty	The penalty to be applied in the regularized likelihood subproblems. 'MCP', and 'lasso' are provided. 'MCP' is recommended.

Value

Output Vector consisting of Rsq mediated(Rsq.mediated), Lower confidence bound constructed by the asymptotic variance (CI_asym_low), Upper confidence bound constructed by the asymptotic variance (CI_asym_up), Lower confidence bound constructed by the conservative variance (CI_cons_low), Upper confidence bound constructed by the conservative variance (CI_cons_up), number of selected mediators in subsample 1 (pab1), number of selected mediators in subsample 2 (pab2), and the Rsq that used to calculate the Rsq measure: variance of outcome explained by mediator (Rsq.YM), variance of outcome explained by the independent variable (Rsq.YX), and variance of outcome explained by mediator and independent variable (Rsq.YMX); Sample Size in analysis (Sample Size)

Name of selected mediators in subsample 1 (select1)

Name of selected mediators in subsample 2 (select2)

Examples

```
{
  data(example)
  attach(example)
  CF_Rsq.measure(Y=Y, M=M, X=X, tune = "bic", penalty = "MCP")
}
```

CI.Rsq.measure	<i>Functions to generate the confidence interval of the Rsq measure using nonparametric bootstrap.</i>
----------------	--

Description

Functions to generate the confidence interval of the Rsq measure using nonparametric bootstrap.

Usage

```
CI.Rsq.measure(
  p = 1/2,
  Y,
  M,
  Covar = NULL,
  X,
  method = c("iSIS", "ALL"),
  B = 200,
  iter.max = 3,
  nsis = NULL,
  init.FDR.cutoff = 0.1,
  filtering = FALSE
)
```

Arguments

<code>p</code>	Proportion of the training dataset for selecting mediators regarding to the whole dataset, default is set as 1/2
<code>Y</code>	Vector of outcome type of interest; Only Gaussian distributed outcome is accepted.
<code>M</code>	Matrix of putative mediators
<code>Covar</code>	Covariate matrix
<code>X</code>	Vector of the exposure or independent variable of interest, e.g. environmental exposure
<code>method</code>	Method used to screen out non-mediators. When no variable selection is required, <code>method='ALL'</code> ; otherwise, iterative sure independence screening (SIS) is used for variable selection, i.e., <code>method='iSIS'</code> .
<code>B</code>	Number of bootstrap samples, default is 100
<code>iter.max</code>	Maximum number of iteration used in iSIS, default=10 (details see the SIS package)
<code>nsis</code>	Number of predictors recruited by iSIS
<code>init.FDR.cutoff</code>	FDR threshold for the filtering method on M1 type of mediators.
<code>filtering</code>	TRUE: filtering mediators based on the strength of indep and mediators as a pre-processing step; FALSE: all putative mediators are included, default=FALSE.

Value

CI: The 95 percent confidence intervals of Rsq measure (Rsq.mediated), shared over simple effects (SOS), number of mediators selected (pab), variance of outcome explained by mediator (Rsq.YM), variance of outcome explained by the independent variable (Rsq.YX), and variance of outcome explained by mediator and independent variable (Rsq.YMX). The estimates for each bootstrap are also returned.

Examples

```
{
  data(example)
  attach(example)
  CI.Rsq.measure( p=1/2, Y=Y,M=M,X=X,method='ALL', B=1, iter.max=1)
}
```

example

*Example dataset***Description**

Example dataset for RsqMed.

Format

Contains the following objects:

Y a vector of 100 observations, representing outcome.

X a vector of 100 observations, representing independent variable.

M a 100 * 100 matrix, representing mediators.

Cov a 100 * 2 matrix, representing covariates.

Rsq.measure

Function to calculate the Rsq function as a total mediation effect size measure (Gaussian outcome only). If method = 'iSIS', a two-step procedure is performed, where the first step filters the non-mediators based on the first p proportion of the data and the second step calculates the point estimates for Rsq using random-effect models on the remaining data. If method = 'ALL', Rsq is calculated based on all subjects and mediators (assuming all mediators are the true mediators). It is optional to adding filtering step on putative mediators to exclude M1 type of non-mediators (See Yang et al, BMC bioinformatics, 2021).

Description

Function to calculate the Rsq function as a total mediation effect size measure (Gaussian outcome only). If method = 'iSIS', a two-step procedure is performed, where the first step filters the non-mediators based on the first p proportion of the data and the second step calculates the point estimates for Rsq using random-effect models on the remaining data. If method = 'ALL', Rsq is calculated based on all subjects and mediators (assuming all mediators are the true mediators). It is optional to adding filtering step on putative mediators to exclude M1 type of non-mediators (See Yang et al, BMC bioinformatics, 2021).

Usage

```
Rsq.measure(
  p = 1/2,
  Y,
  M,
  Covar = NULL,
```

```

X,
method = c("iSIS", "ALL"),
iter.max = 3,
nsis = NULL,
init.FDR.cutoff = 0.1,
filtering = FALSE
)

```

Arguments

p	Proportion of the training dataset for selecting mediators regarding the whole dataset, default is set as 1/2. If method = 'ALL', all subjects are included.
Y	Vector of outcome; Only Gaussian distributed outcome is accepted.
M	Matrix of putative mediators
Covar	Covariate matrix or vector, default = NULL
X	Vector of the independent variable of interest, e.g. environmental variable
method	Method used to screen out M2 type of non-mediators. When method = 'ALL', no variable selection is performed and Rsq is calculated on all data; otherwise, iterative sure independence screening (SIS) is performed on training dataset, i.e., method = 'iSIS'.
iter.max	Maximum number of iteration used in iSIS, default = 3
nsis	Number of variables recruited by iterative SIS (please refer the SIS package for detailed explanations)
init.FDR.cutoff	FDR threshold for the filtering after SIS.
filtering	When filtering = TRUE, filtering based on the strength of independent variable and mediators after SIS is performed; When filtering = FALSE, no preprocessing is performed before variable selection. By default filtering = FALSE.

Value

Output Vector consists of Rsq mediated (Rsq.mediated), shared over simple effects (SOS), number of selected mediators (pab), and the Rsq that used to calculate the Rsq measure: variance of outcome explained by mediator (Rsq.YM), variance of outcome explained by the independent variable (Rsq.YX), and variance of outcome explained by mediator and independent variable (Rsq.YMX), n.train is the sample size on which variable selection is performed, n.estimate is the sample size based on which the mediation effect is estimated.

Name of selected putative mediators (select). Note that M1 type of non-mediators may still be included in the model, but it would not impact the estimation of total mediation effect under certain assumptions.

Examples

```

{
## Not run:
data(example)

```

```
attach(example)
Rsq.measure(p=1/2, Y=Y, M=M,Covar=Cov,X=X,method='iSIS', iter.max=1, filtering=TRUE)

## End(Not run)
}
```

Index

CF_Rsq.measure, [2](#)

CI.Rsq.measure, [3](#)

example, [5](#)

Rsq.measure, [5](#)