

# Package ‘robustMVMR’

May 2, 2020

**Title** Perform the Robust Multivariable Mendelian Randomization Analysis

**Version** 0.1.0

**Date** 2020-04-28

**Author** Zhao Yang

**Maintainer** Zhao Yang <yangz98@connect.hku.hk>

**Description** Perform the robust multivariable Mendelian randomization 'robustMVMR' analysis in the two-sample Mendelian randomization setting. The 'robustMVMR' package produces both the robust estimators and the robust standard errors via the MM-estimates, which has been demonstrated to protect against heteroskedasticity, autocorrelation, and the presence of outliers in Yohai (1987) <doi:10.1214/aos/1176350366> and Croux (2003) <https://EconPapers.repec.org/RePEc:ete:ceswps:ces0316>.

**Depends** R (>= 3.5.0)

**Imports** lmtest (>= 0.9-37), robustbase (>= 0.93-5), ggplot2 (>= 3.2.1), stats (>= 3.6.2)

**License** GPL-2 | GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.0

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2020-05-02 13:20:02 UTC

## R topics documented:

IL1_LUSC . . . . .	2
robustMVMR . . . . .	2

<b>Index</b>	<b>6</b>
--------------	----------

---

 IL1\_LUSC

*IL-1 and squamous cell lung cancer*


---

### Description

Data from two genome-wide association studies for proteomics and the internal lung cancer consortium study. The outcome is the squamous cell lung cancer, and the exposures are interleukin-1 alpha, interleukin-1 beta, and interleukin-1 receptor antagonist.

### Usage

```
data(IL1_LUSC)
```

### Format

An matrix contains the genetic associations of the exposures and the outcome.

### References

- 1. Sun et al. (2018) Genomic atlas of the human plasma proteome. **Nature**, 558(7708):73-79
- 2. Ahola-Olli et al. (2017) Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. **Am J Hum Genet**, 100(1):40-50.
- 3. Wang et al. (2014) Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. **Nat Genet**, 46(7):736-41.

---

 robustMVMR

*Perform the robust multivariable mendelian randomization analysis*


---

### Description

The robustMVMR perform the robust multivariable Mendelian randomization (robustMVMR) analysis in the two-sample MR setting based on the MM-estimator. The conventional multivariable Mendelian randomization (MVMR) estimate the causal effect by employing the weighted least square estimators, in which the inverse variance of the SNPs-outcome association is arbitrarily selected as the weights with an additional assumption about the heteroskedastic error. When all the instrument assumptions of MVMR are satisfied; that is,

- the variant is associated with at least 1 of the risk factor;
- the variant is not associated with a confounder of any of the risk-outcome associations;
- the variant is conditionally independent of the outcome given the risk factors and confounders;
- the variants are required to be independent;
- the heteroskedastic error;
- the linearity and homogeneity of all associations.

Violation of any one of the aforementioned assumptions can cause severe bias in MVMR. The robustMVMR produces the robust causal effect and robust standard errors based on the MM-estimates, which has been demonstrated to protect against the heteroskedasticity, autocorrelation, and the presence of outliers. The interested reader is referred to Yohai (1987) paper and Crousx et al (2004) paper. In MR setting, outliers of the multi-instruments may indicate the horizontal pleiotropic effect, which has been comprehensively discussed in [www.nature.com/articles/s41588-018-0099-7](http://www.nature.com/articles/s41588-018-0099-7) Verbank et al (2018) paper in univariable MR setting.

Notable, the assumption of heteroskedastic error in MVMR setting has a key role in estimating the causal effect and its standard error. However, recent advances suggest that the estimated effect derived from the conventional MVMR may be biased, especially when the exposures are highly correlated and the correlation matrix of these exposures is unknown. And, such scenarios often happen in the MVMR setting, especially when the multiple independent instruments are clumped by using the 1000 Genome Project as the reference. Furthermore, the function of these selected variants are not fully understood; that is, the horizontal pleiotropy (or outliers) may also arise. The results from Verbank et al (2018) paper reported that almost half (around 48%) of significant causal relationship in MR suffered from the horizontal pleiotropy. The effect of the horizontal pleiotropy on the "true" causal estimates ranged from -131% to 201%, with a false-positive rate of 10%. In such a case, the robustMVMR would provide better estimates and standard errors of the causal effect than those in the conventional MVMR.

Furthermore, the overall conditional F-statistic for testing conditional weak instrument bias and the modified Q-statistic for testing the instrument validity in the multivariable Mendelian randomization proposed by Sanderson et al. (2020) are also provided in **robustMVMR**. Along with the overall conditional F-statistic, a pairwise conditional F-statistic matrix is also provided to identify the possible source of conditional weak instrument bias.

Lastly, the data-driven result about the correlation matrix of exposure is also reported in **robustMVMR**. Such a matrix is derived from the standard error of each exposure under the heteroskedasticity assumption. The further using of this matrix should be cautious.

## Usage

```
robustMVMR(
  betaGY,
  sebetaGY,
  pvalbetaGY,
  betaGX,
  sebetaGX,
  pvalbetaGX,
  pval_threshold = 1e-05,
  plot = FALSE
)
```

## Arguments

betaGY	A numeric vector of the beta-coefficient for the SNPs-outcome associations. For the binary outcome, the log-odds ratio estimates from the logistic regression analysis are strongly recommended.
sebetaGY	The numeric vector of the standard errors for the SNPs-outcome associations.

pvalbetaGY	The numeric vector of P values for the SNPs-outcome associations.
betaGX	A matrix of the beta-coefficient for the SNPs-exposures associations.
sebetaGX	The matrix of the standard error for the SNPs-exposure associations.
pvalbetaGX	The matrix of P values for the SNPs-exposure associations.
pval_threshold	The threshold of the P value for selecting the genetic variants for exposures. By default, pval_threshold = 1e-05.
plot	The option for return the scatter plot with the marginal effect of each exposure. By default, plot = FALSE.

### Value

A list contains nine components, including

**Breusch\_Pagen\_test** The **Breusch Pagen test** for the heteroskedasticity assumption. Rejecting the NULL hypothesis indicate the violation of the heteroskedastic error.

**mvMRResult\_homo\_robust** The results from the robust multivariable Mendelian randomization with the weights being 1.

**mvMRResult\_heter\_burgess** The results from the conventional MVMR analysis with the weights being  $1/sebetaGY^2$ .

**mvMRResult\_heter\_robust** The results from the robust multivariable Mendelian randomization with the weights being  $1/sebetaGY^2$ . Of these, the conditional F-statistic is also reported.

**marginalEffect** The results from the robust univariable Mendelian randomization based on the validity instruments with the weights being  $1/sebetaGY^2$ .

**Conditional\_F\_statistic\_matrix** The pair-wise conditional F-statistics of exposures included in the robust MVMR analysis.

**Q\_pleiotropy\_test** The modified Q-statistic for testing the instrument validity used in the robust MVMR analysis.

**rho\_Exposures** The correlation matrix of exposures included in the robust MVMR analysis. **It is worth noting that this is a data-driven result.**

**plots** The scatter plot of the marginal effect of each exposure on the outcome.

### References

- 1. Yohai, V.J. (1987) High breakdown-point and high efficiency robust estimates for regression. **The Annals of Statistics**, pp.642-656.
- 2. Croux, C., Dhaene, G. and Hoorelbeke, D. (2004) Robust standard errors for robust estimators. **CES-Discussion paper series (DPS) 03.16**, pp.1-20.
- 3. Verbanck, M., Chen, C.Y., Neale, B. and Do, R. (2018) Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. **Nature genetics**, 50(5), pp.693-698.
- 4. Sanderson, E., Davey Smith, G., Windmeijer, F. and Bowden, J. (2019) An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. **International journal of epidemiology**, 48(3), pp.713-727.

**See Also**

[bptest](#) and [lmrob](#)

**Examples**

```
data(IL1_LUSC)
## -- SNP-outcome data
betaGY <- IL1_LUSC[, "beta.LUSC"]
sebetaGY <- IL1_LUSC[, "sebeta.LUSC"]
pvalbetaGY <- IL1_LUSC[, "pval.LUSC"]
## -- SNP-exposure data
betaGX <- IL1_LUSC[, c("beta.IL1A_Sun", "beta.IL1B_Ahola", "beta.IL1RA_Ahola")]
sebetaGX <- IL1_LUSC[, c("se.IL1A_Sun", "se.IL1B_Ahola", "se.IL1RA_Ahola")]
pvalbetaGX <- IL1_LUSC[, c("pval.IL1A_Sun", "pval.IL1B_Ahola", "pval.IL1RA_Ahola")]
## -- Robust MVMR
fit <- robustMVMR(betaGY = betaGY, sebetaGY = sebetaGY, pvalbetaGY = pvalbetaGY,
                 betaGX = betaGX, sebetaGX = sebetaGX, pvalbetaGX = pvalbetaGX,
                 pval_threshold = 1e-05, plot = FALSE)
## -- Main results of the robust MVMR
fit$mvMRRResult_heter_robust
## -- The modified Q-statistic for testing instrument validity
fit$Q_pleiotropy_test
## -- The pair-wise conditional F-statistic matrix
fit$Conditional_F_statistic_matrix
## -- The correlation matrix of the exposures
fit$rho_Exposures
```

# Index

\*Topic **datasets**

IL1\_LUSC, [2](#)

bptest, [5](#)

IL1\_LUSC, [2](#)

lmrob, [5](#)

robustMVMR, [2](#)