Package: MVMR (via r-universe)

August 27, 2024

Title MVMR

Version 0.4

Description An R package for performing multivariable Mendelian randomization analyses.

License GPL-3

URL https://github.com/WSpiller/MVMR, https://wspiller.github.io/MVMR/, https://mrcieu.r-universe.dev/MVMR

BugReports https://github.com/WSpiller/MVMR/issues

Depends R (>= 3.6)

Imports boot

Suggests knitr, MendelianRandomization, rmarkdown

VignetteBuilder knitr

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Repository https://mrcieu.r-universe.dev

RemoteUrl https://github.com/WSpiller/MVMR

RemoteRef HEAD

RemoteSha 65705da9421b6235c7458dba6f01cddfebfe96f5

Contents

`format_mvmr` *format_mvmr*

Description

Reads in summary data. Checks and organises columns for use in calculating multivariable Mendelian Randomization analyses. Where variant IDs are not provided, a vector is generated for variant identification.

Usage

```
format_mvmr(BXGs, BYG, seBXGs, seBYG, RSID)
```

Arguments

`BXGs`	A matrix containing beta-coefficient values for genetic associations with the each exposure. Columns should indicate exposure number, with rows representing estimates for a given genetic variant.
`BYG`	A numeric vector of beta-coefficient values for genetic associations with the outcome.
`seBXGs`	A matrix containing standard errors corresponding to the matrix of beta-coefficients `BXGs`.
`seBYG`	A numeric vector of standard errors corresponding to the beta-coefficients `BYG`.
`RSID`	A vector of names for genetic variants included in the analysis. If variant IDs are not provided (`RSID="NULL"`), a vector of ID numbers will be generated.

Value

A formatted data frame of class `mvmr_format`.

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
r_input <- format_mvmr(
    BXGs = rawdat_mvmr[,c("LDL_beta","HDL_beta")],
    BYG = rawdat_mvmr$SBP_beta,
    seBXGs = rawdat_mvmr[,c("LDL_se","HDL_se")],
    seBYG = rawdat_mvmr$SBP_se,
    RSID = rawdat_mvmr$SNP)
names(r_input)
class(r_input)
```

`ivw_mvmr` *ivw_mvmr*

Description

Fits an IVW multivariable Mendelian randomization model using first order weights.

Usage

```
ivw_mvmr(r_input, gencov = 0)
```

Arguments

`r_input`	A formatted data frame using the `format_mvmr` function or an object of class `MRMVInput` from `MendelianRandomization::mr_mvinput`
`gencov`	Calculating heterogeneity statistics requires the covariance between the effect of the genetic variants on each exposure to be known. This can either be estimated from individual level data, be assumed to be zero, or fixed at zero using non-overlapping samples of each exposure GWAS. A value of `0` is used by default.

Value

An dataframe containing MVMR results, including estimated coefficients, their standard errors, t-statistics, and corresponding (two-sided) p-values.

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
r_input <- format_mvmr(
    BXGs = rawdat_mvmr[,c("LDL_beta","HDL_beta")],
    BYG = rawdat_mvmr$SBP_beta,
    seBXGs = rawdat_mvmr[,c("LDL_se","HDL_se")],
    seBYG = rawdat_mvmr$SBP_se,
    RSID = rawdat_mvmr$SNP)
ivw_mvmr(r_input)
```

---

mrmvinput_to_mvmr_format

Convert an object of class MRMVInput from the MendelianRandomization package to the MVMR mvmr_format class

---

Description

Creates a data.frame with additional class `mvmr_format` from an object of class `MRMVInput` generated by `MendelianRandomization::mr_mvinput`.

Usage

```
mrmvinput_to_mvmr_format(dat)
```

Arguments

`dat`	Object from `MendelianRandomization::mr_mvinput`.

Value

Object of class `mvmr_format`, the MVMR format

Examples

```
if (require("MendelianRandomization", quietly = TRUE)) {
bx <- as.matrix(rawdat_mvmr[,c("LDL_beta", "HDL_beta")])
bxse <- as.matrix(rawdat_mvmr[,c("LDL_se", "HDL_se")])
dat <- MendelianRandomization::mr_mvinput(bx = bx,
                                          bxse = bxse,
                                          by = rawdat_mvmr$SBP_beta,
                                          byse = rawdat_mvmr$SBP_se,
                                          snps = rawdat_mvmr$SNP)
dat <- mrmvinput_to_mvmr_format(dat)
head(dat)
class(dat)
}
```

mvmr *mvmr (legacy)*

Description

Note: This function is from the old version of the MVMR package and will be replaced in the future: The gencov argument should be set to zero when using `mvmr()`.

Usage

```
mvmr(r_input, gencov, weights)
```

Arguments

`r_input`	A formatted data frame using the `format_mvmr` function or an object of class `MRMVInput` from `MendelianRandomization::mr_mvinput`
`gencov`	Calculating heterogeneity statistics requires the covariance between the effect of the genetic variants on each exposure to be known. This can either be estimated from individual level data, be assumed to be zero, or fixed at zero using non-overlapping samples of each exposure GWAS. A value of `0` is used by default.
`weights`	A value specifying the inverse variance weights used to calculate IVW estimate and Cochran's Q statistic. Currently only first order weights are available (`1`).

Details

Fits an IVW multivariable Mendelian randomization model using first order weights. The function returns an object of class `"MVMRIVW"`, containing regression estimates, estimated heterogeneity as a measure of instrument strength (`Q_strength`), and estimated heterogeneity as a measure of instrument validity (`Q_valid`).

Value

An object of class `"MVMRIVW"` containing the following components:

`summary` A summary of the MVMR regression model, including estimated coefficients, standard errors, t-statistics, p-values, and heterogeneity statistics.

`coef` The estimated coefficients, their standard errors, t-statistics, and corresponding (two-sided) p-values.

`Q_strength` A data frame displaying modified Cochran's Q statistics for assessing instrument strength with respect to each exposure. The Q-statistic increases proportionally with instrument strength, and analogous to univariate MR analyses, a value equal to or greater than 10 can be used as a minimum threshold for instrument strength. Note that for these statistics it is not informative to evaluate p-values.

`Q_valid` A modified form of Cochran's Q statistic measuring heterogeneity in causal effect estimates obtained using each genetic variant. Observed heterogeneity is indicative of a violation of the exclusion restriction assumption in MR (validity), which can result in biased effect estimates.

`p_valid` A p-value corresponding to the heterogeneity measure for instrument validity (`Q_valid`)

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
# Example using format_mvmr formatted data
r_input <- format_mvmr(
    BXGs = rawdat_mvmr[,c("LDL_beta","HDL_beta")],
    BYG = rawdat_mvmr$SBP_beta,
    seBXGs = rawdat_mvmr[,c("LDL_se","HDL_se")],
    seBYG = rawdat_mvmr$SBP_se,
    RSID = rawdat_mvmr$SNP)
mvmr(r_input, 0, 1)

# Example using MRMVInput formatted data from the MendelianRandomization package
if (require("MendelianRandomization", quietly = TRUE)) {
bx <- as.matrix(rawdat_mvmr[,c("LDL_beta", "HDL_beta")])
bxse <- as.matrix(rawdat_mvmr[,c("LDL_se", "HDL_se")])
dat <- MendelianRandomization::mr_mvinput(bx = bx,
                                          bxse = bxse,
                                          by = rawdat_mvmr$SBP_beta,
                                          byse = rawdat_mvmr$SBP_se,
                                          snps = rawdat_mvmr$SNP)
mvmr(r_input = r_input, gencov = 0, weights = 1)
}
```

---

phenocov_mvmr *phenocov_mvmr*

Description

Uses an external phenotypic covariance matrix and summary data to estimate covariance matrices for estimated effects of individual genetic variants on each exposure. The phenotypic covariance matrix should be constructed using standardised phenotype measures. The function returns a number of covariance matrices equal to the number of SNPs, where SNP and row numbers reference ordered exposures.

Usage

```
phenocov_mvmr(Pcov, seBXGs)
```

Arguments

`Pcov`	A phenotypic matrix using exposures, constructed using individual level exposure data. Columns should be ordered by exposure so as to match `format_mvmr`.
`seBXGs`	A matrix containing standard errors corresponding in relation to the gene-exposure association for each SNP.

Value

A list of covariance matrices with respect to each genetic variant, retaining the ordering in `seBXGs`

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
## Not run:
phenocov_mvmr(Pcov, summarydata[,c(3,4)])

## End(Not run)
```

`pleiotropy_mvmr` *pleiotropy_mvmr*

Description

Calculates modified form of Cochran's Q statistic measuring heterogeneity in causal effect estimates obtained using each genetic variant. Observed heterogeneity is indicative of a violation of the exclusion restriction assumption in MR (validity), which can result in biased effect estimates. The function takes a formatted dataframe as an input, obtained using the function `format_mvmr`. Additionally, covariance matrices for estimated effects of individual genetic variants on each exposure can also be provided. These can be estimated using external data by applying the `snpcov_mvmr` or `phenocov_mvmr` functions, are input manually. The function returns a dataframe including the conditional Q-statistic for instrument validity, and a corresponding P-value.

Usage

```
pleiotropy_mvmr(r_input, gencov = 0)
```

Arguments

r_input	A formatted data frame using the format_mvmr function or an object of class MRMVInput from MendelianRandomization::mr_mvinput
gencov	Calculating heterogeneity statistics requires the covariance between the effect of the genetic variants on each exposure to be known. This can either be estimated from individual level data, be assumed to be zero, or fixed at zero using non-overlapping samples of each exposure GWAS. A value of 0 is used by default.

Value

A Q-statistic for instrument validity and the corresponding p-value

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
## Not run:
pleiotropy_mvmr(r_input, covariances)

## End(Not run)
```

---

qhet_mvmr *qhet_mvmr*

Description

Fits a multivariable Mendelian randomization model adjusting for weak instruments. The functions requires a formatted dataframe using the format_mvmr function, as well a phenotypic correlation matrix pcor. This should be obtained from individual level phenotypic data, or constructed as a correlation matrix where correlations have previously been reported. Confidence intervals are calculated using a non-parametric bootstrap. By default, standard errors are not produced but can be calculated by setting se = TRUE. The number of bootstrap iterations is specified using the iterations argument. Note that calculating confidence intervals at present can take a substantial amount of time.

Usage

```
qhet_mvmr(r_input, pcor, CI, iterations)
```

Arguments

r_input	A formatted data frame using the format_mvmr function or an object of class MRMVInput from MendelianRandomization::mr_mvinput
pcor	A phenotypic correlation matrix including the correlation between each exposure included in the MVMR analysis.
CI	Indicates whether 95 percent confidence intervals should be calculated using a non-parametric bootstrap.
iterations	Specifies number of bootstrap iterations for calculating 95 percent confidence intervals.

Value

An dataframe containing effect estimates with respect to each exposure.

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
## Not run:
qhet_mvmr(r_input, pcor, CI = TRUE, iterations = 1000)

## End(Not run)
```

---

rawdat_mvmr — *Raw multivariable MR summary data using lipid fractions as exposures and systolic blood pressure as an outcome.*

Description

A dataset containing summary data on 145 genetic variants associated with either low-density lipoprotein (LDL), high-density lipoprotein (HDL), or triglycerides. Data includes variant rsid numbers, associations with each lipid fraction, the associations between genetic variants and systolic blood pressure, and corresponding standard errors.

Usage

```
rawdat_mvmr
```

Format

A data frame with 145 rows and 9 variables. A full description of the data is available by using the command `vignette("MVMR")`.

Details

rawdat_mvmr

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

Source

- http://www.mrbase.org/
- https://www.nature.com/articles/ng.2797
- https://www.nature.com/articles/ng.3768

---

snpcov_mvmr *snpcov_mvmr*

Description

Uses individual level genetic and exposure data to generate covariance matrices for estimated effects of individual genetic variants on each exposure. The function returns a number of covariance matrices equal to the number of SNPs, where SNP and row numbers reference ordered exposures.

Usage

```
snpcov_mvmr(Gs, Xs)
```

Arguments

`Gs`	A matrix or dataframe containing genetic instrument measures. Columns should indicate genetic variant number, with rows representing an observed measure of the genetic variant.
`Xs`	A matrix or dataframe containing exposure measures. Columns should indicate exposure number, with rows representing an observed measure for the given exposure.

Value

A list of covariance matrices with respect to each genetic variant, retaining the ordering in `Gs`

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
## Not run:
snpcov_mvmr(data[,1:10], data[,11:13])

## End(Not run)
```

---

strength_mvmr *strength_mvmr*

Description

Calculates the conditional F-statistic for assessing instrument strength in two sample summary multivariable Mendelian randomization. The function takes a formatted dataframe as an input, obtained using the function `format_mvmr`. Additionally, covariance matrices for estimated effects of individual genetic variants on each exposure can also be provided. These can be estimated using external data by applying the `snpcov_mvmr` or `phenocov_mvmr` functions, are input manually. The function returns a dataframe including the conditional F-statistic with respect to each exposure. A conventional F-statistic threshold of 10 is used in basic assessments of instrument strength.

Usage

```
strength_mvmr(r_input, gencov = 0)
```

Arguments

`r_input`	r_input A formatted data frame using the `format_mvmr` function or an object of class `MRMVInput` from `MendelianRandomization::mr_mvinput`
`gencov`	Calculating heterogeneity statistics requires the covariance between the effect of the genetic variants on each exposure to be known. This can either be estimated from individual level data, be assumed to be zero, or fixed at zero using non-overlapping samples of each exposure GWAS. A value of `0` is used by default.

Value

A dataframe showing the conditional F-statistic for each exposure.

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2018, 48, 3, 713-727. Available from: doi:10.1093/ije/dyy262

Examples

```
## Not run:
strength_mvmr(data, covariances)

## End(Not run)
```

`strhet_mvmr` *strhet_mvmr*

Description

Calculates the conditional F-statistic for assessing instrument strength in two sample summary multivariable Mendelian randomization through minimisation of Q-statistics. The function takes a formatted dataframe as an input, obtained using the function `format_mvmr`. Additionally, covariance matrices for estimated effects of individual genetic variants on each exposure can also be provided. These can be estimated using external data by applying the `snpcov_mvmr` or `phenocov_mvmr` functions, are input manually. The function returns a dataframe including the conditional F-statistic with respect to each exposure. A conventional F-statistic threshold of 10 is used in basic assessments of instrument strength.

Usage

```
strhet_mvmr(r_input, gencov)
```

Arguments

`r_input`	A formatted data frame using the `format_mvmr` function or an object of class `MRMVInput` from `MendelianRandomization::mr_mvinput`
`gencov`	Calculating heterogeneity statistics requires the covariance between the effect of the genetic variants on each exposure to be known. This can either be estimated from individual level data, be assumed to be zero, or fixed at zero using non-overlapping samples of each exposure GWAS. A value of `0` is used by default.

Value

A dataframe showing the conditional F-statistic for each exposure.

Author(s)

Wes Spiller; Eleanor Sanderson; Jack Bowden.

References

Sanderson, E., et al., An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. International Journal of Epidemiology, 2019, 48, 3, 713-727. doi:10.1093/ije/dyy262

Examples

```
## Not run:
strhet_mvmr(r_input, covariances)

## End(Not run)
```

Index