

Package: ColliderBias (via r-universe)

March 1, 2025

Type Package

Title Adjust collider bias and weak instrument bias in Mendelian Randomisation

Version 1.0.1

Author Frank Dudbridge, Siyang Cai

Maintainer Frank Dudbridge <frank.dudbridge@leicester.ac.uk>, Siyang Cai <sc814@leicester.ac.uk>

Description This package provides adjustment for collider bias and weak instrument bias in association statistics in the context of a genome-wide association study for a subsequent event, using CWLS (Corrected Weighted Least Squares), MR-RAPS and Slope Hunter. An estimation of true causal effect between disease progression and an exposure of interest can be computed using generalised instrument effect regression with CWBLS (Corrected Weighted Bivariate Least Squares).

License GPL-3

URL <https://github.com/SiyangCai/ColliderBias>

BugReports <https://github.com/SiyangCai/ColliderBias/issues/>

Depends R (>= 3.5)

Imports boot, mr.raps, SlopeHunter

Remotes Osmahmoud/SlopeHunter

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Config/pak/sysreqs make libicu-dev libssl-dev

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

RemoteUrl <https://github.com/SiyangCai/ColliderBias>

RemoteRef HEAD

RemoteSha 66606d074c5cfd5263a0e2234ceb0a88fc9fdb1

Contents

CWBLS	2
methodCB	3
testData	5
Index	6

CWBLS	<i>Estimate true causal effect in bivariate Mendelian randomization to adjust for collider bias and weak instrument bias.</i>
-------	---

Description

This function is designed for adjusting collider bias and weak instrument bias to estimate unbiased causal effect between an exposure trait and disease progression trait, using generalised instrumental effect regression and CWBLS adjustment in bivariate Mendelian randomization.

Usage

```
CWBLS(data)
```

Arguments

data	A data frame of input dataset. This dataset should contain following columns: xbeta and xse represents the effects of exposure trait and its standard errors, dbeta and dse represents the effects of disease trait and its standard errors, ybeta and yse represents the effects of disease progression trait and its standard errors.
------	---

Value

An object of class "CWBLS" which contains:

- b The true causal estimate between exposure and disease progression
- b.se Standard error of b

Author(s)

Siyang Cai, Frank Dudbridge

References

Cai S, Hartley A, Mahmoud O, Tilling K, Dudbridge F. Adjusting for collider bias in genetic association studies using instrumental variable methods. *Genet Epidemiol.* 2022;46:303–16.

Examples

```
# Load the test dataset
data(testData)

# Find the true causal between exposure and disease progression.
CWBS(testData)
```

methodCB	<i>Adjust association statistics for collider bias and weak instrument bias in Mendelian randomization.</i>
----------	---

Description

Given effect sizes and standard errors for predictors of an index trait and a subsequent trait, this function adjusts the statistics for the subsequent trait for selection bias through the index trait.

Usage

```
methodCB(
  xbeta,
  xse,
  ybeta,
  yse,
  prune = NULL,
  method,
  weighted = "1",
  model = "ivw",
  od = FALSE
)
```

Arguments

xbeta	Vector of effects on the index trait.
xse	Vector of standard errors of xbeta.
ybeta	Vector of effects on the subsequent trait.
yse	Vector of standard errors of ybeta.
prune	Vector containing the indices of an approximately independent subset of the predictors in xbeta and ybeta. If unspecified, all predictors will be used.
method	Method to adjust for weak instrument bias in the regression of ybeta[prune] on xbeta[prune]. "CWLS" applies Corrected Weighted Least Squares estimator, "SH" applies Slope Hunter and "mr.raps" applies MR-RAPS.
weighted	Weights to be used in raw regression and CWLS. Default is using "1" as the first order weight, while "0" indicates unweighted model and "2" applies second order weight.
model	Regression model to be used in raw regression and CWLS . Default is "IVW" and could be replaced by "Egger" to use MR-Egger model.
od	Whether to use "over.dispersion" in MR-RAPS.

Details

Effect sizes are on a linear scale, so could be the coefficients from linear regression, or log odds ratios, or log hazard ratios. Effects on the subsequent trait are regressed on the effects on the index trait. By default, the regression is weighted by the inverse variances of the subsequent trait effects, so users are recommended to use the default `weighted = "1"`. Other weight options are developed for any potential application from the user. The regression coefficient is firstly adjusted for weak instrument bias through the method, and the residuals are then used to obtain adjusted effect sizes and standard errors for the subsequent trait to correct for collider bias. Such an approach is called instrumental effect regression.

The regression should be performed on a subset of predictors that are independent. In the context of a genome-wide association study, these would be LD-pruned SNPs. In terms of the input parameters, the regression command is `lm(ybeta[prune]~xbeta[prune], weights=1/yse[prune]^2)`.

The effects in `xbeta` and `ybeta` should be aligned for the same variables and the same direction prior to running the adjustment using this package.

Value

An object of class "methodCB" which contains:

- `ybeta.adj` Adjusted effects on the subsequent trait
- `yse.adj` Adjusted standard errors of `ybeta.adj`
- `ychisq.adj` Chi-square statistics for $(ybeta.adj/yse.adj)^2$
- `yp.adj` P-values for `ychisq.adj` on 1df
- `b` Coefficient of the regression of `ybeta[prune]` on `xbeta[prune]`, after correction for regression dilution
- `b.se` Standard error of `b`
- `b.ci` Lower and upper confidence limits for `b`
- `b.raw` Regression coefficient without correction for regression dilution
- `b.raw.se` Standard error of `b.raw`

Author(s)

Siyang Cai, Frank Dudbridge

References

Cai S, Hartley A, Mahmoud O, Tilling K, Dudbridge F. Adjusting for collider bias in genetic association studies using instrumental variable methods. *Genet Epidemiol.* 2022;46:303–16.

Mahmoud O, Dudbridge F, Smith G D, et al. Slope-Hunter: A robust method for index-event bias correction in genome-wide association studies of subsequent traits[J]. *bioRxiv*, 2020.

Zhao, Q, et al. "Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score." *The Annals of Statistics* 48.3 (2020): 1742-1769.

Examples

```
# Load the test dataset
data(testData)

# Adjust for collider bias using instrumental effect regression,
# and weak instrument bias using CWLS.
methodCB(testData$dbeta, testData$dse, testData$ybeta, testData$yse, method = "CWLS")
```

testData	<i>Sample simulated genome-wide association study (GWAS) summary statistics for analysis.</i>
----------	---

Description

A simple `data_frame` that contains genetic summary data with three traits for example showcases, An exposure X, a disease incidence trait D and a disease progression trait Y. Data contains 10,000 independent SNPs under Hardy-Weinberg equilibrium with minor allele frequencies drawn uniformly from (0.01, 0.49). All SNPs had independent effects on both X and D and also directly on the outcome Y to create invalid instruments. SNP effects, confounders and residual variation in exposure X and disease trait D were then drawn independently from normal distributions, assuming no genetic effects between disease trait and exposure. Variance explained by non-genetic confounders is fixed as 20%. The heritability takes values of 30% for all three traits. The true causal effect between the exposure of interest and the disease progression (outcome) is 0.4.

Usage

```
testData
```

Format

A `data_frame` with six columns

xbeta A list of SNP-exposure effects given a random exposure of our interest.

xse Standard errors of SNP-exposure effects.

dbeta A list of SNP-disease effects.

dse Standard errors of SNP-disease effects.

ybeta A list of SNP-outcome effects.

yse Standard errors of SNP-outcome effects.

Index

* **datasets**
 testData, 5

CWBLs, 2

methodCB, 3

testData, 5