Package 'GCIM'

February 27, 2025

Title The genetic causality inference model(GCIM) is a statistical method for detecting the causal direction in GxE interaction studies

Version 0.0.1.000

Description What the package does (one paragraph).

License `use_mit_license()`, `use_gpl3_license()` or friends to pick a license

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2.9000

R topics documented:

`bp_dis_cov` *A phenotype data from the discovery dataset containing binary phenotype trated as exposure for the proposed causal directions.*

Description

This dataset contains covariate information for individuals in the discovery dataset, including various confounders for a binary outcome.

Usage

```
data(bp_dis_cov)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary Covariate A binary covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

`bp_dis_phen` *A phenotype data from the discovery dataset containing binary phenotype trated as an outcome for the proposed causal directions.*

Description

This dataset contains phenotype for the outcome information for individuals in the discovery dataset, for a binary outcome treated as a binary outcome for the proposed causal diections.

Usage

```
data(bp_dis_phen)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary outcome A binary outcome (1=controls, 2=cases)

bp_tar_cov *Covariate Data File for Discovery Dataset with Binary phenotype*

Description

This dataset contains covariate information for individuals in the target dataset, including various confounders for a binary phenotype for the outcome.

Usage

```
data(bp_tar_cov)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary Covariate A binary covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

bp_tar_phen *Covariate Data File for target Dataset with Binary Outcomes*

Description

This dataset contains phenotype for the outcome information for individuals in the target dataset, including various confounders for a binary outcome.

Usage

```
data(bp_tar_phen)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary outcome A binary outcome (0=controls, 1=cases)

br_dis_cov *A phenotype data file from the discovery dataset, containing the binary exposure treated as binary outcomes for the reverse direction.*

Description

This dataset contains phenotype for the outcome information for individuals in the discovery dataset, for a binary exposure as a binary outcome in the reverse direction.

Usage

```
data(br_dis_cov)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary outcome A binary outcome (1=controls, 2=cases)

br_dis_phen *A phenotype data file from the discovery dataset, containing the binary oucome treated as binary exposure for the reverse direction.*

Description

This dataset contains phenotype information for individuals in the discovery dataset, for a binary outcome treated as a binary exposure in the reverse direction.

Usage

```
data(br_dis_phen)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary Covariate A binary covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

br_tar_cov *A phenotype data file from the target dataset, containing the binary exposure treated as an oucome for the reverse direction of causation*

Description

This dataset contains phenotype data for the outcome information for individuals in the target dataset, for a binary outcome as a binary outcome in the reverse directions of causations.

Usage

```
data(br_tar_cov)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary outcome A binary outcome (0=controls, 1=cases)

br_tar_phen *A phenotype data file from target dataset, containing a binary outcome treated as a binary exposure in the reverse direction of causations*

Description

This dataset contains phenotypic data for the covariate information for individuals in the target dataset, for a binary phenotype for the outcome, which is treated as a binary exposure and covariates .

Usage

```
data(br_tar_phen)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

Binary Covariate A binary covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

b_gwas *Perform GWAS for covariates.*

Description

Perform GWAS for covariates.

Usage

```
b_gwas(plink_path, dis_snp, bp_dis_cov)
```

Arguments

plink_path	Path to the PLINK executable application.
dis_snp	Prefix for binary files for the discovery dataset.
bp_dis_cov	File path for covariate data in the discovery dataset.
temp_dir	Directory to save output files.

Value

None. Results are saved to files.

b_gweis *Perform GWEIS for binary outcome and exposure variables.*

Description

This function performs genome-wide-by-environemnt interaction studies (GWEIS) and processes the results to generate files for downstream analysis.

Usage

```
b_gweis(plink_path, dis_snp, bp_dis_phen, bp_dis_cov)
```

Arguments

plink_path	Path to the PLINK executable application.
dis_snp	Prefix for binary files for the discovery dataset.
bp_dis_phen	File path for the phenotype data in the discovery dataset.
bp_dis_cov	File path for covariate data in the discovery dataset.
temp_dir	Directory to save output files.

Value

None. Results are saved to files.

dis_snp.bim *PLINK .bim File for SNP Information*

Description

This dataset follows the standard PLINK .bim format and contains SNP-related information for the discovery dataset, including chromosome ID, SNP ID, and allele details.

Usage

```
data(dis_snp.bim)
```

Format

A data frame with 6 columns:

Column 1 Chromosome ID (integer)

Column 2 SNP ID (character)

Column 3 Position in centimorgans (numeric)

Column 4 Base-pair coordinate (integer)

Column 5 Minor Allele (character)

Column 6 Reference Allele (character)

dis_snp.fam *PLINK .fam File for Family and Phenotype Information*

Description

This dataset follows the standard PLINK .fam format and contains family structure and phenotype information for individuals in the discovery dataset.

Usage

```
data(dis_snp.fam)
```

Format

A data frame with 6 columns:

Column 1 Family ID (character or integer)

Column 2 Individual ID (character or integer)

Column 3 Father's ID (character or integer; 0 if missing)

Column 4 Mother's ID (character or integer; 0 if missing)

Column 5 Sex (integer; 1 = male, 2 = female, 0 = unknown)

Column 6 Phenotype value (numeric; 1 = control, 2 = case, -9 = missing)

dis_snp.map *PLINK .map File for SNP Information*

Description

This dataset follows the standard PLINK .map file format and contains genetic marker information, including chromosome, SNP ID, position in centimorgans, and base-pair coordinates.

Usage

```
data(dis_snp.map)
```

Format

A data frame with 4 columns:

Column 1 Chromosome ID (integer)

Column 2 SNP ID (character)

Column 3 Position in centimorgans (numeric; typically 0 if unknown)

Column 4 Base-pair coordinate (integer; position on the chromosome in base pairs)

dis_snp.ped *PLINK .ped File for Genotype and Sample Information*

Description

This dataset follows the standard PLINK .ped file format. It contains family, individual, parental, sex, phenotype, and genotype information for each sample.

Usage

```
data(dis_snp.ped)
```

Format

A data frame with 6 required columns, followed by genotype data:

Column 1 Family ID (character)

Column 2 Individual ID (character)

Column 3 Paternal ID (0 if unknown, character)

Column 4 Maternal ID (0 if unknown, character)

Column 5 Sex (1 = male, 2 = female, 0 = unknown)

Column 6 Phenotype value (1 = control, 2 = case, -9 = missing)

Columns 7+ Genotype data in pairs of alleles for each SNP (e.g., A T, G G, coded as separate columns).

gcim_b *Perform Regression Analysis for GCIM with binary outcome.*

Description

Perform Regression Analysis for GCIM with binary outcome.

Usage

```
gcim_b(bp_tar_phen, bp_tar_cov, Add_PRS, Int_PRS, Cov_PRS, confounders)
```

Arguments

bp_tar_phen	File path for the target phenotype data.
bp_tar_cov	File path for the target covariate data.
Additive	Scaled additive PRS values.
Interaction	Scaled interaction PRS values.
Covariate	Scaled covariate PRS values.
Confounders	Data frame of additional confounders.

Value

Summary of the regression model.

gcim_q *Perform Regression Analysis for GCIM with quantitative outcome.*

Description

Perform Regression Analysis for GCIM with quantitative outcome.

Usage

```
gcim_q(qp_tar_phen, qp_tar_cov, Add_PRS, Int_PRS, Cov_PRS, confounders)
```

Arguments

bp_tar_phen	File path for the target phenotype data.
bp_tar_cov	File path for the target covariate data.
Additive	Scaled additive PRS values.
Interaction	Scaled interaction PRS values.
Covariate	Scaled covariate PRS values.
Confounders	Data frame of additional confounders.

Value

Summary of the regression model.

prs_scores *Compute Polygenic Risk Scores (PRS).*

Description

Compute Polygenic Risk Scores (PRS).

Usage

```
prs_scores(plink_path, tar_snp)
```

Arguments

plink_path	Path to the PLINK executable application.
tar_snp	Prefix for binary files for the target dataset.
temp_dir	Directory to save output files.

Value

A list containing scaled PRS values for additive, interaction, and covariate scores.

qp_dis_cov *Covariate Data File for Discovery Dataset with quantitative phenotype*

Description

This dataset contains covariate information for individuals in the discovery dataset, including various confounders for a quantitative covariate.

Usage

```
data(qp_dis_cov)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative Covariate A quantitative covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

qp_dis_phen *Covariate Data File for Discovery Dataset with Binary Outcomes*

Description

This dataset contains phenotype for the outcome information for individuals in the discovery dataset, including various confounders for a quantitative outcome.

Usage

```
data(qp_dis_phen)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative outcome A quantitative outcome #'

qp_tar_cov *Covariate Data File for Discovery Dataset with Binary phenotype*

Description

This dataset contains covariate information for individuals in the target dataset, including various confounders for a quantitative phenotype for the outcome.

Usage

```
data(qp_tar_cov)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative Covariate A quantitative covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

qp_tar_phen *Covariate Data File for target Dataset with Binary Outcomes*

Description

This dataset contains phenotype for the outcome information for individuals in the target dataset, including various confounders for a quantitative outcome.

Usage

```
data(qp_tar_phen)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative outcome A qunatitative outcome #'

qr_dis_cov *Covariate data File for discovery dataset with quantitative exposure as an outcomes for the reverse direction.*

Description

This dataset contains phenotype for the outcome information for individuals in the discovery dataset, including various confounders for a quantitative outcome.

Usage

```
data(qr_dis_cov)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative outcome A quantitative outcome #'

qr_dis_phen *Covariate data File for discovery dataset of the outcome as quantitative covariate phenotype for the reverse direction.*

Description

This dataset contains covariate information for individuals in the discovery dataset, including various confounders and the outcome phenotype as a quantitative covariate.

Usage

```
data(qr_dis_phen)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative Covariate A quantitative covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

qr_tar_cov *Covariate data file for target dataset of the covariate with quantitative outcomes for the reverse direction of causation*

Description

This dataset contains phenotype for the outcome information for individuals in the target dataset, including various confounders as a quantitative outcome of the covariate.

Usage

```
data(qr_tar_cov)
```

Format

A data frame with 800 rows and 3 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative outcome A qunatitative outcome #'

qr_tar_phen — *Covariate data file for discovery dataset of the covariate with quantitative phenotype for the reverse*

Description

This dataset contains covariate information for individuals in the target dataset, including various confounders for a quantitative phenotype of the quantiative ciovariate for the reverse direction of causation.

Usage

```
data(qr_tar_phen)
```

Format

A data frame with 800 rows and 19 columns:

Family ID Family identification number

Individual ID Individual identification number

quantitative Covariate A quantitative covariate

Confounder 1 Confounder 1 description ...

Confounder 16 Confounder 16 description

q_gwas — *Perform GWAS for covariates.*

Description

Perform GWAS for covariates.

Usage

```
q_gwas(plink_path, dis_snp, qp_dis_cov)
```

Arguments

plink_path	Path to the PLINK executable application.
dis_snp	Prefix for binary files for the discovery dataset.
qp_dis_cov	File path for covariate data in the discovery dataset.
temp_dir	Directory to save output files.

Value

None. Results are saved to files.

q_gweis — *Perform GWEIS for quantitative outcome and quantitative exposure variables.*

Description

This function performs genome-wide interaction studies (GWEIS) and processes the results to generate files for downstream analysis.

Usage

```
q_gweis(plink_path, dis_snp, qp_dis_phen, qp_dis_cov)
```

Arguments

plink_path	Path to the PLINK executable application.
dis_snp	Prefix for binary files for the discovery dataset.
qp_dis_phen	File path for the phenotype data in the discovery dataset.
qp_dis_cov	File path for covariate data in the discovery dataset.
temp_dir	Directory to save output files.

Value

None. Results are saved to files.

tar_snp.bim — *PLINK .bim file for SNP information*

Description

This dataset follows the standard PLINK .bim format and contains SNP-related information for the target dataset, including chromosome ID, SNP ID, and allele details.

Usage

```
data(tar_snp.bim)
```

Format

A data frame with 6 columns:

Column 1 Chromosome ID (integer)

Column 2 SNP ID (character)

Column 3 Position in centimorgans (numeric)

Column 4 Base-pair coordinate (integer)

Column 5 Minor Allele (character)

Column 6 Reference Allele (character)

tar_snp.fam *PLINK .fam file for family and phenotype information*

Description

This dataset follows the standard PLINK .fam format and contains family structure and phenotype information for individuals in the target dataset.

Usage

```
data(tar_snp.fam)
```

Format

A data frame with 6 columns:

Column 1 Family ID (character or integer)

Column 2 Individual ID (character or integer)

Column 3 Father's ID (character or integer; 0 if missing)

Column 4 Mother's ID (character or integer; 0 if missing)

Column 5 Sex (integer; 1 = male, 2 = female, 0 = unknown)

Column 6 Phenotype value (numeric; 1 = control, 2 = case, -9 = missing)

tar_snp.map *PLINK .map file for SNP information*

Description

This dataset follows the standard PLINK .map file format and contains genetic marker information, including chromosome, SNP ID, position in centimorgans, and base-pair coordinates.

Usage

```
data(tar_snp.map)
```

Format

A data frame with 4 columns:

Column 1 Chromosome ID (integer)

Column 2 SNP ID (character)

Column 3 Position in centimorgans (numeric; typically 0 if unknown)

Column 4 Base-pair coordinate (integer; position on the chromosome in base pairs)

`tar_snp.ped` *PLINK .ped file for genotype and sample information*

Description

This dataset follows the standard PLINK .ped file format. It contains family, individual, parental, sex, phenotype, and genotype information for each sample.

Usage

```
data(tar_snp.ped)
```

Format

A data frame with 6 required columns, followed by genotype data:

Column 1 Family ID (character)

Column 2 Individual ID (character)

Column 3 Paternal ID (0 if unknown, character)

Column 4 Maternal ID (0 if unknown, character)

Column 5 Sex (1 = male, 2 = female, 0 = unknown)

Column 6 Phenotype value (1 = control, 2 = case, -9 = missing)

Columns 7+ Genotype data in pairs of alleles for each SNP (e.g., A T, C, G, coded as separate columns).

Index