Package 'GxEprs'

July 26, 2023

Title Genotype-by-
environment (GxE) Interaction in Polygenic Risk Score Models for Quantitative and Binary Traits

Version 1.0

Description
A novel PRS model is introduced to enhance the prediction accuracy by utilising GxE effects.
This package performs Genome Wide Association Studies (GWAS) and Genome Wide Environment Interaction Studies (GWEIS)
using a discovery dataset. The package has the ability to obtain polygenic risk scores (PRSs) for a target sample.
Finally it predicts the risk values of each individual in the target sample. Users have the choice of using existing
models (Li et al., 2015)<doi.org/10.1093/annonc/mdu565>, (Pandis et al., 2013)<doi.org/10.1093/ejo/cjt054>,
(Peyrot et al., 2018)<doi.org/10.1016/j.biopsych.2017.09.009> and (Song et al., 2022)<doi.org/10.1038/s41467-022-32407-9>,
as well as newly proposed models for genomic risk prediction (refer to the URL for more details).

URL https://github.com/DoviniJ/GxEprs

License GPL (>=3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

Depends R (>= 2.10)

LazyData true

R topics documented:

`Bcov_discovery` *Covariate data file of the discovery dataset when the outcome is binary This contains covariate information of the individuals in the discovery dataset following confounders*

Description

Covariate data file of the discovery dataset when the outcome is binary This contains covariate information of the individuals in the discovery dataset following confounders

Usage

```
Bcov_discovery
```

Format

A dataframe with 7916 rows and 18 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Standardized covariate

Column 4 Square of the standardized covariate

Column 5 Confounder 1

Column 6 Confounder 2

Column 7 Confounder 3

Column 8 Confounder 4

Column 9 Confounder 5

Column 10 Confounder 6

Column 11 Confounder 7

Column 12 Confounder 8

Column 13 Confounder 9

Column 14 Confounder 10

Column 15 Confounder 11

Column 16 Confounder 12

Column 17 Confounder 13

Column 18 Confounder 14

`Bcov_target` *Covariate data file of the target dataset when the outcome is binary This contains covariate information of the individuals in the target dataset following confounders*

Description

Covariate data file of the target dataset when the outcome is binary This contains covariate information of the individuals in the target dataset following confounders

Usage

```
Bcov_target
```

Format

A dataframe with 1939 rows and 18 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Standardized covariate

Column 4 Square of the standardized covariate

Column 5 Confounder 1

Column 6 Confounder 2

Column 7 Confounder 3

Column 8 Confounder 4

Column 9 Confounder 5

Column 10 Confounder 6

Column 11 Confounder 7

Column 12 Confounder 8

Column 13 Confounder 9

Column 14 Confounder 10

Column 15 Confounder 11

Column 16 Confounder 12

Column 17 Confounder 13

Column 18 Confounder 14

`Bphe_discovery` *Phenotype data file of the discovery dataset when the outcome is binary This contains phenotype information of the individuals in the discovery dataset*

Description

Phenotype data file of the discovery dataset when the outcome is binary This contains phenotype information of the individuals in the discovery dataset

Usage

```
Bphe_discovery
```

Format

A dataframe with 7916 rows and 3 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Phenotype (1=controls, 2=cases)

`Bphe_target` *Phenotype data file of the target dataset when the outcome is binary This contains phenotype information of the individuals in the target dataset*

Description

Phenotype data file of the target dataset when the outcome is binary This contains phenotype information of the individuals in the target dataset

Usage

```
Bphe_target
```

Format

A dataframe with 1939 rows and 3 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Phenotype (0=controls, 1=cases)

`DummyData.bim` *PLINK .bim file*

Description

PLINK .bim file

Usage

```
DummyData.bim
```

Format

This follows PLINK general format

Column 1 Chromosome ID

Column 2 SNP ID

Column 3 Position of centimorgans

Column 4 Base-pair coordinate

Column 5 Minor Allele

Column 6 Reference Allele

`DummyData.fam` *PLINK .fam file*

Description

PLINK .fam file

Usage

```
DummyData.fam
```

Format

This follows PLINK general format

Column 1 Family ID

Column 2 Individual ID

Column 3 Father's ID

Column 4 Mother's ID

Column 5 Sex

Column 6 Phenotype value

DummyData.map *PLINK .map file*

Description

PLINK .map file

Usage

```
DummyData.map
```

Format

This follows PLINK general format

DummyData.ped *PLINK .ped file*

Description

PLINK .ped file

Usage

```
DummyData.ped
```

Format

This follows PLINK general format

GWAS_binary *GWAS_binary function This function performs GWAS using plink2 and outputs the GWAS summary statistics with additive SNP effects. Users may save the output in a user-specified file (see example).*

Description

GWAS_binary function This function performs GWAS using plink2 and outputs the GWAS summary statistics with additive SNP effects. Users may save the output in a user-specified file (see example).

Usage

```
GWAS_binary(plink_path, b_file, Bphe_discovery, Bcov_discovery, thread = 20)
```

Arguments

plink_path	Path to the PLINK executable application
b_file	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
Bphe_discovery	Name (with file extension) of the phenotype file containing family ID, individual ID and phenotype of the discovery dataset as columns, without heading
Bcov_discovery	Name (with file extension) of the covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the discovery dataset as columns, without heading
thread	Number of threads used

Value

This function will perform GWAS and output

B_out.trd.sum	GWAS summary statistics with additive SNP effects

Examples

```
## Not run:
x <- GWAS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery,
thread = 20)
sink("B_out.trd.sum") #to create a file in the working directory
write.table(x, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x) #to obtain the head of GWAS summary statistics of additive SNP effects
x$V1 #to extract the chromosome number (CHROM)
x$V2 #to extract the base pair position (POS)
x$V3 #to extract the SNP ID (ID)
x$V4 #to extract the reference allele (REF)
x$V5 #to extract the alternate allele (ALT)
x$V6 #to extract the minor allele (A1)
x$V7 #to extract whether firth regression is used (FIRTH?)
x$V8 #to extract the type of test performed (TEST)
x$V9 #to extract the number of allele observations (OBS_CT)
x$V10 #to extract the odds ration of the SNP effect (OR)
x$V11 #to extract the standard error of log odds (LOG(OR)_SE)
x$V12 #to extract the test statistic (Z_STAT)
x$V13 #to extract the p value (P)
x$V14 #to extract the error code (ERRCODE)

## End(Not run)
```

GWAS_quantitative *GWAS_quantitative function This function performs GWAS using plink2 and outputs the GWAS summary statistics with additive SNP effects. Users may save the output in a user-specified file (see example).*

Description

GWAS_quantitative function This function performs GWAS using plink2 and outputs the GWAS summary statistics with additive SNP effects. Users may save the output in a user-specified file (see example).

Usage

```
GWAS_quantitative(
  plink_path,
  b_file,
  Qphe_discovery,
  Qcov_discovery,
  thread = 20
)
```

Arguments

plink_path	Path to the PLINK executable application
b_file	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
Qphe_discovery	Name (with file extension) of the phenotype file containing family ID, individual ID and phenotype of the discovery dataset as columns, without heading
Qcov_discovery	Name (with file extension) of the covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the discovery dataset as columns, without heading
thread	Number of threads used

Value

This function will perform GWAS and output

Q_out.trd.sum	GWAS summary statistics with additive SNP effects

Examples

```
## Not run:
x <- GWAS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery,
thread = 20)
sink("Q_out.trd.sum") #to create a file in the working directory
write.table(x, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x) #to obtain the head of GWAS summary statistics of additive SNP effects
x$V1 #to extract the chromosome number (CHROM)
x$V2 #to extract the base pair position (POS)
x$V3 #to extract the SNP ID (ID)
x$V4 #to extract the reference allele (REF)
x$V5 #to extract the alternate allele (ALT)
x$V6 #to extract the minor allele (A1)
x$V7 #to extract whether firth regression is used (FIRTH?)
x$V8 #to extract the type of test performed (TEST)
x$V9 #to extract the number of allele observations (OQS_CT)
```

```
x$V10 #to extract the odds ration of the SNP effect (OR)
x$V11 #to extract the standard error of log odds (LOG(OR)_SE)
x$V12 #to extract the test statistic (Z_STAT)
x$V13 #to extract the p value (P)
x$V14 #to extract the error code (ERRCODE)

## End(Not run)
```

`GWEIS_binary` *GWEIS_binary function This function performs GWEIS using plink2 and outputs the GWEIS summary statistics with additive SNP effects and interaction SNP effects separately. It is recommended to save the outputs in separate user-specified files (see examples).*

Description

GWEIS_binary function This function performs GWEIS using plink2 and outputs the GWEIS summary statistics with additive SNP effects and interaction SNP effects separately. It is recommended to save the outputs in separate user-specified files (see examples).

Usage

```
GWEIS_binary(plink_path, b_file, Bphe_discovery, Bcov_discovery, thread = 20)
```

Arguments

`plink_path`	Path to the PLINK executable application
`b_file`	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
`Bphe_discovery`	Phenotype file containing family ID, individual ID and phenotype of the discovery dataset as columns, without heading
`Bcov_discovery`	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the discovery dataset as columns, without heading
`thread`	Number of threads used

Value

This function will perform GWEIS and output

`B_out.add.sum`	GWEIS summary statistics with additive SNP effects
`B_out.gxe.sum`	GWEIS summary statistics with interaction SNP effects

Examples

```
## Not run:
x <- GWEIS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery,
thread = 20)
sink("B_out.add.sum") #to create a file in the working directory
write.table(x[[1]], sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
sink("B_out.gxe.sum") #to create a file in the working directory
write.table(x[[2]], sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x[[1]]) #to extract the head of all columns in GWEIS summary statistics of
additive SNP effects
x[[1]]$V1 #to extract the chromosome number (CHROM)
x[[1]]$V2 #to extract the base pair position (POS)
x[[1]]$V3 #to extract the SNP ID (ID)
x[[1]]$V4 #to extract the reference allele (REF)
x[[1]]$V5 #to extract the alternate allele (ALT)
x[[1]]$V6 #to extract the minor allele (A1)
x[[1]]$V7 #to extract whether firth regression is used (FIRTH?)
x[[1]]$V8 #to extract the type of test performed (TEST)
x[[1]]$V9 #to extract the nmber of allele observations (OBS_CT)
x[[1]]$V10 #to extract the odds ration of the SNP effect (OR)
x[[1]]$V11 #to extract the standard error of log odds (LOG(OR)_SE)
x[[1]]$V12 #to extract the test statistic (Z_STAT)
x[[1]]$V13 #to extract the p value (P)
x[[1]]$V14 #to extract the error code (ERRCODE)
head(x[[2]]) #to extract the head of all columns in GWEIS summary statistics of
interaction SNP effects
x[[2]]$V1 #to extract the chromosome number (CHROM)
x[[2]]$V2 #to extract the base pair position (POS)
x[[2]]$V3 #to extract the SNP ID (ID)
x[[2]]$V4 #to extract the reference allele (REF)
x[[2]]$V5 #to extract the alternate allele (ALT)
x[[2]]$V6 #to extract the minor allele (A1)
x[[2]]$V7 #to extract whether firth regression is used (FIRTH?)
x[[2]]$V8 #to extract the type of test performed (TEST)
x[[2]]$V9 #to extract the number of allele observations (OBS_CT)
x[[2]]$V10 #to extract the odds ration of the SNP effect (OR)
x[[2]]$V11 #to extract the standard error of log odds (LOG(OR)_SE)
x[[2]]$V12 #to extract the test statistic (Z_STAT)
x[[2]]$V13 #to extract the p value (P)
x[[2]]$V14 #to extract the error code (ERRCODE)

## End(Not run)
```

`GWEIS_quantitative` *GWEIS_quantitative function This function performs GWEIS using plink2 and outputs the GWEIS summary statistics with additive SNP effects and interaction SNP effects separately. It is recommended to save the outputs in separate user-specified files (see examples).*

Description

GWEIS_quantitative function This function performs GWEIS using plink2 and outputs the GWEIS summary statistics with additive SNP effects and interaction SNP effects separately. It is recommended to save the outputs in separate user-specified files (see examples).

Usage

```
GWEIS_quantitative(
  plink_path,
  b_file,
  Qphe_discovery,
  Qcov_discovery,
  thread = 20
)
```

Arguments

plink_path	Path to the PLINK executable application
b_file	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
Qphe_discovery	Phenotype file containing family ID, individual ID and phenotype of the discovery dataset as columns, without heading
Qcov_discovery	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the discovery dataset as columns, without heading
thread	Number of threads used

Value

This function will perform GWEIS and output

Q_out.add.sum	GWEIS summary statistics with additive SNP effects
Q_out.gxe.sum	GWEIS summary statistics with interaction SNP effects

Examples

```
## Not run:
x <- GWEIS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery,
thread = 20)
sink("Q_out.add.sum") #to create a file in the working directory
write.table(x[[1]], sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
sink("Q_out.gxe.sum") #to create a file in the working directory
write.table(x[[2]], sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x[[1]]) #to read the head of all columns in GWEIS summary statistics of
additive SNP effects
x[[1]]$V1 #to extract the chromosome number (CHROM)
x[[1]]$V2 #to extract the base pair position (POS)
```

```
x[[1]]$V3 #to extract the SNP ID (ID)
x[[1]]$V4 #to extract the reference allele (REF)
x[[1]]$V5 #to extract the alternate allele (ALT)
x[[1]]$V6 #to extract the minor allele (A1)
x[[1]]$V7 #to extract the type of test performed (TEST)
x[[1]]$V8 #to extract the nmber of allele observations (OBS_CT)
x[[1]]$V9 #to extract the SNP effect (BETA)
x[[1]]$V10 #to extract the standard error of each SNP effect (SE)
x[[1]]$V11 #to extract the test statistic (T_STAT)
x[[1]]$V12 #to extract the p value (P)
x[[1]]$V13 #to extract the error code (ERRCODE)
head(x[[2]]) #to read the head of all columns in GWEIS summary statistics of
interaction SNP effects
x[[2]]$V1 #to extract the chromosome number (CHROM)
x[[2]]$V2 #to extract the base pair position (POS)
x[[2]]$V3 #to extract the SNP ID (ID)
x[[2]]$V4 #to extract the reference allele (REF)
x[[2]]$V5 #to extract the alternate allele (ALT)
x[[2]]$V6 #to extract the minor allele (A1)
x[[2]]$V7 #to extract the type of test performed (TEST)
x[[2]]$V8 #to extract the number of allele observations (OBS_CT)
x[[2]]$V9 #to extract the SNP effect (BETA)
x[[2]]$V10 #to extract the standard error of each SNP effect (SE)
x[[2]]$V11 #to extract the test statistic (T_STAT)
x[[2]]$V12 #to extract the p value (P)
x[[2]]$V13 #to extract the error code (ERRCODE)

## End(Not run)
```

`PRS_binary` *PRS_binary function This function uses plink2 and outputs Polygenic Risk Scores (PRSs) of all the individuals, using pre-generated GWAS and/or GWEIS summary statistics. Note that the input used in this function can be generated by using GWAS_binary and/or GWEIS_binary functions. Users may save the output in a user-specified file (see examples).*

Description

PRS_binary function This function uses plink2 and outputs Polygenic Risk Scores (PRSs) of all the individuals, using pre-generated GWAS and/or GWEIS summary statistics. Note that the input used in this function can be generated by using GWAS_binary and/or GWEIS_binary functions. Users may save the output in a user-specified file (see examples).

Usage

```
PRS_binary(plink_path, b_file, summary_input)
```

Arguments

`plink_path`	Path to the PLINK executable application
`b_file`	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
`summary_input`	Pre-generated GWAS and/or GWEIS summary statistics

Value

This function will output

`B_trd.sscore` PRSs for each individual

Examples

```
## Not run:
a <- GWAS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery)
b <- GWEIS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery)
x <- PRS_binary(plink_path, DummyData, summary_input = a)
sink("B_trd.sscore") #to create a file in the working directory
write.table(x, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x) #to read the head of all columns in the output
x$V1 #to extract the family ID's of target dataset (FID)
x$V2 #to extract the individual ID's of target dataset (IID)
x$V3 #to extract the number of alleles across scored variants (ALLELE_CT)
x$V4 #to extract the sum of named allele dosages (NAMED_ALLELE_DOSAGE_SUM)
x$V5 #to extract the polygenic risk scores (SCORE1_AVG)
y <- PRS_binary(plink_path, DummyData, summary_input = b[[1]])
sink("B_add.sscore") #to create a file in the working directory
write.table(y, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
z <- PRS_binary(plink_path, DummyData, summary_input = b[[2]])
sink("B_gxe.sscore") #to create a file in the working directory
write.table(z, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output

## End(Not run)
```

`PRS_quantitative` *PRS_quantitative function This function uses plink2 and outputs Polygenic Risk Scores (PRSs) of all the individuals, using pre-generated GWAS and/or GWEIS summary statistics. Note that the input used in this function can be generated by using GWAS_quantitative and/or GWEIS_quantitative functions. Users may save the output in a user-specified file (see examples).*

Description

PRS_quantitative function This function uses plink2 and outputs Polygenic Risk Scores (PRSs) of all the individuals, using pre-generated GWAS and/or GWEIS summary statistics. Note that the input used in this function can be generated by using GWAS_quantitative and/or GWEIS_quantitative functions. Users may save the output in a user-specified file (see examples).

Usage

```
PRS_quantitative(plink_path, b_file, summary_input)
```

Arguments

plink_path	Path to the PLINK executable application
b_file	Prefix of the binary files, where all .fam, .bed and .bim files have a common prefix
summary_input	Pre-generated GWAS and/or GWEIS summary statistics

Value

This function will output

Q_trd.sscore	PRSs for each individual

Examples

```
## Not run:
a <- GWAS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery)
b <- GWEIS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery)
x <- PRS_quantitative(plink_path, DummyData, summary_input = a)
sink("Q_trd.sscore") #to create a file in the working directory
write.table(x, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
head(x) #to read the head of all columns in the output
x$V1 #to extract the family ID's of target dataset (FID)
x$V2 #to extract the individual ID's of target dataset (IID)
x$V3 #to extract the number of alleles across scored variants (ALLELE_CT)
x$V4 #to extract the sum of named allele dosages (NAMED_ALLELE_DOSAGE_SUM)
x$V5 #to extract the polygenic risk scores (SCORE1_AVG)
y <- PRS_quantitative(plink_path, DummyData, summary_input = b[[1]])
sink("Q_add.sscore") #to create a file in the working directory
write.table(y, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output
z <- PRS_quantitative(plink_path, DummyData, summary_input = b[[2]])
sink("Q_gxe.sscore") #to create a file in the working directory
write.table(z, sep = " ", row.names = FALSE, quote = FALSE) #to write the output
sink() #to save the output

## End(Not run)
```

Qcov_discovery *Covariate data file of the discovery dataset when the outcome is quantitative This contains covariate information of the individuals in the discovery dataset following confounders*

Description

Covariate data file of the discovery dataset when the outcome is quantitative This contains covariate information of the individuals in the discovery dataset following confounders

Usage

```
Qcov_discovery
```

Format

A dataframe with 6426 rows and 18 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Standardized covariate

Column 4 Square of the standardized covariate

Column 5 Confounder 1

Column 6 Confounder 2

Column 7 Confounder 3

Column 8 Confounder 4

Column 9 Confounder 5

Column 10 Confounder 6

Column 11 Confounder 7

Column 12 Confounder 8

Column 13 Confounder 9

Column 14 Confounder 10

Column 15 Confounder 11

Column 16 Confounder 12

Column 17 Confounder 13

Column 18 Confounder 14

`Qcov_target` *Covariate data file of the target dataset when the outcome is quantitative This contains covariate information of the individuals in the target dataset following confounders*

Description

Covariate data file of the target dataset when the outcome is quantitative This contains covariate information of the individuals in the target dataset following confounders

Usage

```
Qcov_target
```

Format

A dataframe with 1579 rows and 18 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Standardized covariate

Column 4 Square of the standardized covariate

Column 5 Confounder 1

Column 6 Confounder 2

Column 7 Confounder 3

Column 8 Confounder 4

Column 9 Confounder 5

Column 10 Confounder 6

Column 11 Confounder 7

Column 12 Confounder 8

Column 13 Confounder 9

Column 14 Confounder 10

Column 15 Confounder 11

Column 16 Confounder 12

Column 17 Confounder 13

Column 18 Confounder 14

`Qphe_discovery` *Phenotype data file of the discovery dataset when the outcome is quantitative This contains phenotype information of the individuals in the discovery dataset*

Description

Phenotype data file of the discovery dataset when the outcome is quantitative This contains phenotype information of the individuals in the discovery dataset

Usage

```
Qphe_discovery
```

Format

A dataframe with 6426 rows and 3 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Phenotype

`Qphe_target` *Phenotype data file of the target dataset when the outcome is quantitative This contains phenotype information of the individuals in the target dataset*

Description

Phenotype data file of the target dataset when the outcome is quantitative This contains phenotype information of the individuals in the target dataset

Usage

```
Qphe_target
```

Format

A dataframe with 1579 rows and 3 columns

Column 1 Family ID

Column 2 Individual ID

Column 3 Phenotype

`summary_permuted_binary` *summary_permuted_binary function This function outputs the p value of permuted model in the target dataset, using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative function. It is recommended to run this function, if you choose to fit 'PRS_gxe x E' interaction component (i.e. novel proposed model, Model 5) when generating risk scores. If the 'PRS_gxe x E' term is significant in Model 5, and insignificant in Model 5* (permuted p value), consider that the 'PRS_gxe x E' interaction component is actually insignificant (always give priority to the p value obtained from the permuted model).*

Description

summary_permuted_binary function This function outputs the p value of permuted model in the target dataset, using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative function. It is recommended to run this function, if you choose to fit 'PRS_gxe x E' interaction component (i.e. novel proposed model, Model 5) when generating risk scores. If the 'PRS_gxe x E' term is significant in Model 5, and insignificant in Model 5* (permuted p value), consider that the 'PRS_gxe x E' interaction component is actually insignificant (always give priority to the p value obtained from the permuted model).

Usage

```
summary_permuted_binary(
  Bphe_target,
  Bcov_target,
  iterations = 1000,
  add_score,
  gxe_score
)
```

Arguments

Bphe_target	Phenotype file containing family ID, individual ID and phenotype of the target dataset as columns, without heading
Bcov_target	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the target dataset as columns, without heading
iterations	Number of iterations used in permutation
add_score	PRSs generated using additive SNP effects of GWEIS summary statistics
gxe_score	PRSs generated using interaction SNP effects of GWEIS summary statistics

Value

This function will output

B_permuted_p.txt
: the p value of the permuted model

Examples

```
## Not run:
a <- GWEIS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery)
p <- PRS_binary(plink_path, DummyData, summary_input = a[[1]])
q <- PRS_binary(plink_path, DummyData, summary_input = a[[2]])
x <- summary_permuted_binary(Bphe_target, Bcov_target, iterations = 1000,
add_score = p, gxe_score = q)
x

## End(Not run)
```

---

summary_permuted_quantitative

summary_permuted_quantitative function This function outputs the p value of permuted model in the target dataset, using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative functions. It is recommended to run this function, if you choose to fit 'PRS_gxe x E' interaction component (i.e. novel proposed model, Model 4) when generating risk scores. If the 'PRS_gxe x E' term is significant in Model 4, and insignificant in Model 4 (permuted p value), consider that the 'PRS_gxe x E' interaction component is actually insignificant (always give priority to the p value obtained from the permuted model).*

Description

summary_permuted_quantitative function This function outputs the p value of permuted model in the target dataset, using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative functions. It is recommended to run this function, if you choose to fit 'PRS_gxe x E' interaction component (i.e. novel proposed model, Model 4) when generating risk scores. If the 'PRS_gxe x E' term is significant in Model 4, and insignificant in Model 4* (permuted p value), consider that the 'PRS_gxe x E' interaction component is actually insignificant (always give priority to the p value obtained from the permuted model).

Usage

```
summary_permuted_quantitative(
  Qphe_target,
  Qcov_target,
  iterations = 1000,
  add_score,
  gxe_score
)
```

Arguments

Qphe_target	Phenotype file containing family ID, individual ID and phenotype of the target dataset as columns, without heading
Qcov_target	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the target dataset as columns, without heading
iterations	Number of iterations used in permutation
add_score	PRSs generated using additive SNP effects of GWEIS summary statistics
gxe_score	PRSs generated using interaction SNP effects of GWEIS summary statistics

Value

This function will output

Q_permuted_p.txt
: the p value of the permuted model

Examples

```
## Not run:
a <- GWEIS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery)
p <- PRS_quantitative(plink_path, DummyData, summary_input = a[[1]])
q <- PRS_quantitative(plink_path, DummyData, summary_input = a[[2]])
x <- summary_permuted_quantitative(Qphe_target, Qcov_target, iterations = 1000,
add_score = p, gxe_score = q)
x

## End(Not run)
```

summary_regular_binary

summary_regular_binary function This function outputs the summary of regular model and final risk score values of each individual in the target dataset using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_binary function.

Description

summary_regular_binary function This function outputs the summary of regular model and final risk score values of each individual in the target dataset using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_binary function.

Usage

```
summary_regular_binary(
  Bphe_target,
  Bcov_target,
  trd_score = NULL,
  add_score = NULL,
  gxe_score = NULL,
  Model
)
```

Arguments

Bphe_target	Phenotype file containing family ID, individual ID and phenotype of the target dataset as columns, without heading
Bcov_target	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the target dataset as columns, without heading
trd_score	PRSs generated using additive SNP effects of GWAS summary statistics
add_score	PRSs generated using additive SNP effects of GWEIS summary statistics
gxe_score	PRSs generated using interaction SNP effects of GWEIS summary statistics
Model	Specify the model number (1: y = PRS_trd + E + PRS_trd x E + confounders, 2: y = PRS_add + E + PRS_add x E + confounders, 3: y = PRS_add + E + PRS_gxe x E + confounders, 4: y = PRS_add + E + PRS_gxe + PRS_gxe x E + confounders, 5: y = PRS_add + E + E^2 + PRS_gxe + PRS_gxe x E + confounders, where y is the outcome variable, E is the covariate of interest, PRS_trd and PRS_add are the polygenic risk scores computed using additive SNP effects of GWAS and GWEIS summary statistics respectively, and PRS_gxe is the polygenic risk scores computed using GxE interaction SNP effects of GWEIS summary statistics.)

Value

This function will output

`Bsummary.txt` the summary of the fitted model

`Individual_risk_values.txt`
the estimated risk values of individuals in the target sample

Examples

```
## Not run:
a <- GWAS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery)
b <- GWEIS_binary(plink_path, DummyData, Bphe_discovery, Bcov_discovery)
p <- PRS_binary(plink_path, DummyData, summary_input = a)
q <- PRS_binary(plink_path, DummyData, summary_input = b[[1]])
r <- PRS_binary(plink_path, DummyData, summary_input = b[[2]])
summary_regular_binary(Bphe_target, Bcov_target,
                            trd_score = p,
                            Model = 1)
summary_regular_binary(Bphe_target, Bcov_target,
                            add_score = q,
                            Model = 2)
summary_regular_binary(Bphe_target, Bcov_target,
                            add_score = q,
                            gxe_score = r,
                            Model = 3)
summary_regular_binary(Bphe_target, Bcov_target,
                            add_score = q,
                            gxe_score = r,
                            Model = 4)
x <- summary_regular_binary(Bphe_target, Bcov_target,
                            add_score = q,
                            gxe_score = r,
                            Model = 5)
sink("Bsummary.txt") #to create a file in the working directory
print(x[[1]][[1]]) #to write the output
sink() #to save the output
sink("Individual_risk_values.txt") #to create a file in the working directory
write.table(x[[2]], sep = " ", row.names = FALSE, col.names = FALSE,
quote = FALSE) #to write the output
sink() #to save the output
x[[1]][[1]] #to obtain the model summary output
x[[1]][[2]] #to extract "Call" of the model summary
x[[1]][[3]] #to extract terms of the model summary
x[[1]][[4]] #to extract family information of the model summary
x[[1]][[5]] #to extract deviance information of the model summary
x[[1]][[6]] #to extract AIC information of the model summary
x[[1]][[7]] #to extract contrasts of the model summary
x[[1]][[8]] #to extract degrees of freedom (df) of residuals of
            #the model summary
x[[1]][[9]] #to extract "Null deviance" of the model summary
x[[1]][[10]] #to extract degrees of freedom (df) of null deviance
             #of the model summary
x[[1]][[11]] #to extract "iter" of the model summary
x[[1]][[12]] #to extract deviance residuals
x[[1]][[13]] #to extract regrerssion coefficients of the model summary
x[[1]][[14]] #to extract aliesed information of the model summary
x[[1]][[15]] #to extract dispersion information of the model summary
x[[1]][[16]] #to extract degrees of freedom of the model summary
x[[1]][[17]] #to extract unscaled variance covariance matrix of all variables
x[[1]][[18]] #to extract scaled variance covariance matrix of all variables
```

```
head(x[[2]]) #to view the head of the predicted risk values of target individuals
x[[2]][,1] #to extract the column containing family ID's
x[[2]][,2] #to extract the column containing individual ID's
x[[2]][,3] #to extract the column containing predicted risk scores

## End(Not run)
```

`summary_regular_quantitative`

summary_regular_quantitative function This function outputs the summary of regular model and final risk score values of each individual in the target dataset using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative function.

Description

summary_regular_quantitative function This function outputs the summary of regular model and final risk score values of each individual in the target dataset using pre-generated Polygenic Risk Scores (PRSs) of all the individuals. Note that the input used in this function can be generated by using PRS_quantitative function.

Usage

```
summary_regular_quantitative(
  Qphe_target,
  Qcov_target,
  trd_score = NULL,
  add_score = NULL,
  gxe_score = NULL,
  Model
)
```

Arguments

`Qphe_target`	Phenotype file containing family ID, individual ID and phenotype of the target dataset as columns, without heading
`Qcov_target`	Covariate file containing family ID, individual ID, standardized covariate, square of standardized covariate, and/or confounders of the target dataset as columns, without heading
`trd_score`	PRSs generated using additive SNP effects of GWAS summary statistics
`add_score`	PRSs generated using additive SNP effects of GWEIS summary statistics
`gxe_score`	PRSs generated using interaction SNP effects of GWEIS summary statistics
`Model`	Specify the model number (1: y = PRS_trd + E + PRS_trd x E + confounders, 2: y = PRS_add + E + PRS_add x E + confounders, 3: y = PRS_add + E + PRS_gxe x E + confounders, 4: y = PRS_add + E + PRS_gxe + PRS_gxe x E + confounders, where y is the outcome variable, E is the covariate of interest, PRS_trd and PRS_add are the polygenic risk scores computed using additive SNP effects of GWAS and GWEIS summary statistics respectively, and PRS_gxe is the polygenic risk scores computed using GxE interaction SNP effects of GWEIS summary statistics.)

Value

This function will output

`Qsummary.txt` the summary of the fitted model

`Individual_risk_values.txt`
the estimated risk values of individuals in the target sample

Examples

```
## Not run:
a <- GWAS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery)
b <- GWEIS_quantitative(plink_path, DummyData, Qphe_discovery, Qcov_discovery)
p <- PRS_quantitative(plink_path, DummyData, summary_input = a)
q <- PRS_quantitative(plink_path, DummyData, summary_input = b[[1]])
r <- PRS_quantitative(plink_path, DummyData, summary_input = b[[2]])
summary_regular_quantitative(Qphe_target, Qcov_target,
                             trd_score = p,
                             Model = 1)
summary_regular_quantitative(Qphe_target, Qcov_target,
                             add_score = q,
                             Model = 2)
summary_regular_quantitative(Qphe_target, Qcov_target,
                             add_score = q,
                             gxe_score = r,
                             Model = 3)
x <- summary_regular_quantitative(Qphe_target, Qcov_target,
                             add_score = q,
                             gxe_score = r,
                             Model = 4)
sink("Qsummary.txt") #to create a file in the working directory
print(x[[1]][[1]]) #to write the output
sink() #to save the output
sink("Individual_risk_values.txt") #to create a file in the working directory
write.table(x[[2]], sep = " ", row.names = FALSE, col.names = FALSE,
quote = FALSE) #to write the output
sink() #to save the output
x[[1]][[1]] #to obtain the model summary output
x[[1]][[2]] #to extract "Call" of the model summary
x[[1]][[3]] #to extract terms of the model summary
x[[1]][[4]] #to extract the residuals
x[[1]][[5]] #to extract regrerssion coefficients of the model summary
x[[1]][[6]] #to extract aliesed information of the model summary
x[[1]][[7]] #to extract "sigma" (residual standard error) information
            #of the model summary
x[[1]][[8]] #to extract degrees of freedom of the model summary
x[[1]][[9]] #to extract the R squared value of the model summary
x[[1]][[10]] #to extract the adjusted R squared value of the model summary
x[[1]][[11]] #to extract the test statistic values of the model summary
x[[1]][[12]] #to extract unscaled variance covariance matrix of all variables
head(x[[2]]) #to view the head of the predicted risk values of target individuals
x[[2]][,1] #to extract the column containing family ID's
x[[2]][,2] #to extract the column containing individual ID's
x[[2]][,3] #to extract the column containing predicted risk scores

## End(Not run)
```

Index