Package 'RobustSNP'

January 1, 2011

Type Package

Title Robust SNP association tests under different genetic models, allowing for covariates

Version 1.0

Depends mvtnorm,car,snpMatrix

Date 2010-07-11

Author Hon-Cheong So

Maintainer Hon-Cheong So<hcso85@gmail.com>

Description This function performs robust genetic association tests. SNP associations are tested under additive, dominant and recessive models. Valid p-values are returned for each SNP, taking into account the multiple testing under different models. Binary or quantitative outcomes are allowed. Covariates may also be included. The program can be applied to genome-wide association studies (GWAS).

License GPL(>=2)

LazyLoad yes

References Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

R topics documented:

`RobustSNP-package` *Robust SNP association tests under different genetic models, allowing for covariates*

Description

This function performs robust genetic association tests. SNP associations are tested under additive, dominant and recessive models. Valid p-values are returned for each SNP, taking into account the multiple testing under different models. Binary or quantitative outcomes are allowed. Covariates may also be included. The program can be applied to genome-wide association studies (GWAS).

Details

Package:	RobustSNP
Type:	Package
Version:	1.0
Depends:	mvtnorm,car,snpMatrix
Date:	2010-07-02
License:	GPL(>=2)
LazyLoad:	yes

Author(s)

Hon-Cheong So Maintainer: Hon-Cheong So <hcso85@gmail.com>

References

Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

`Rbin` *Robust SNP association tests for binary outcomes allowing for covariates*

Description

This function performs robust genetic association tests for binary outcomes. SNP associations are tested under additive, dominant and recessive models. Valid p-values are returned for each SNP, taking into account the multiple testing under different inheritance models. Covariates may also be included.

Usage

```
Rbin(data = NA, outcome = NA, SNP = NA, InputFile = NA, covPres = TRUE,
COVAR = NA, CovarFile = NA, inputheader = TRUE, covarheader = TRUE,
maxpts = 25000, abseps = 0.001)
```

Arguments

`data`	A matrix or dataframe specifying the genotypes. Each row represents an individual and each column represents a SNP. SNPs should be coded as 0,1 or 2 according to the allelic counts.
`outcome`	Binary outcome coded as 0 or 1.
`SNP`	SNP labels. If not specified, SNPs will be labelled by numbers.
`InputFile`	An optional txt file specifying the outcome in the first column and the genotypes in the other columns. Each row represents an individual. SNPs should be coded as 0,1 or 2 according to the allelic counts. Users may either specify both `data` and `outcome` or just specify `InputFile`.
`covPres`	Whether covariates are present. If `covPres` is `TRUE`, the user must specify either `COVAR` OR `CovarFile`.
`COVAR`	(Optional) A matrix or dataframe specifying the covariates. Each row represents an individual and each column represents a covariate.
`CovarFile`	(Optional) A file location for covariates. The file should be readable by the `read.table` command in R.
`inputheader`	(Optional) Whether `InputFile` contains headers. Only applies when `InputFile` is given.
`covarheader`	(Optional) Whether `CovarFile` contains headers. Only applies when `CovarFile` is given.
`maxpts`	(Optional) Arguments for `pmvnorm` in the `mvtnorm` package. Specifies the maximum number of function values (as integer).
`abseps`	(Optional) Arguments for `pmvnorm` in the `mvtnorm` package. Specifies the absolute error tolerance (as double).

Details

The p-values are calculated based on score tests. We compute the covariance matrix of the test statistics under the three genetic models and correct for multiple testing via trivariate integration. Please see the reference (So and Sham 2010) for further details.

Users may use the –recodeA command from PLINK to convert ped files to the desired input format.

Value

The function returns a dataframe with 9 columns. The columns shown include:

`SNP`	The SNP labels
`Z.add`	Z statistic under the additive model
`Z.rec`	Z statistic under the recessive model
`Z.dom`	Z statistic under the dominant model
`P.add`	P-value under the additive model
`P.rec`	P-value under the recessive model
`P.dom`	P-value under the dominant model
`theoP`	The "combined" p-value for the SNP, adjusted for multiple testing under different genetic models
`integ.error`	Estimated absolute error from trivariate integration

Author(s)

Hon-Cheong So

References

Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

See Also

Rbin.block

Examples

```
## Simuate a case-control study with 300 cases/300 controls
## and 10 SNPs
set.seed(3)
outcome1 = c( rep(0,300), rep(1,300) )
data1 = matrix(nrow=600, ncol=10)
for (k in 1:10) {
data1[,k]=sample ( rep(0:2,200) ) }
COVAR1= rnorm(600,mean=100,sd=30)

## with covariates
Rbin( data= data1, outcome=outcome1, SNP=1:10,covPres=TRUE,
COVAR= COVAR1)

##without covariates
Rbin( data= data1, outcome=outcome1, SNP=1:10,covPres=FALSE)
```

Rbin.block — *Robust SNP association tests for binary outcomes allowing for covariates, with genotypes read in blocks*

Description

This function is similar to Rbin, but the genotypes are read in blocks (instead of reading all of them simultaneously). The aim to to reduce the memory consumption. This function is particularly suited for large-scale association studies like genome-wide association studies (GWAS).

Usage

```
Rbin.block(GenoFile = NA, covPres = TRUE, COVAR = NA, CovarFile = NA,
SNPrange = NA, covarheader = TRUE, blocksize = 5000, maxpts = 25000,
abseps = 0.001)
```

Arguments

GenoFile	A file location for genotype file in PLINK binary PED format. A collection of three files with the same filename but different suffices(bed, bim and fam) should be present. Only the filename but NOT the suffix should be given.
covPres	Whether covariates are present. If covPres is TRUE, the user must specify either COVAR OR CovarFile.
COVAR	(Optional) A matrix or dataframe specifying the covariates. Each row represents an individual and each column represents a covariate.
CovarFile	(Optional) A file location for covariates. The file should be readable by the read.table command in R.
SNPrange	(Optional) A numeric vector specifying the range of SNPs to be tested.
covarheader	(Optional) Whether CovarFile contains headers. Only applies when CovarFile is given.
blocksize	The number of SNPs to be read at one time.
maxpts	(Optional) Arguments for pmvnorm in the mvtnorm package. Specifies the maximum number of function values (as integer).
abseps	(Optional) Arguments for pmvnorm in the mvtnorm package. Specifies the absolute error tolerance (as double).

Details

This function reduces memory consumption by reading a limited number of SNPs at one time. The block size should be adjusted based on the size of the dataset and the size of the usable memory. Decreasing the block size will reduce memory consumption but may prolong the analysis due to increased time in loading the data.

Value

Same as the ouput from Rbin.

Author(s)

Hon-Cheong So

References

Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

See Also

Rbin

Examples

```
## Assume the binary PED file is named "gwas" and the covariate file is "covar.txt"
## Result = Rbin.block (GenoFile="gwas" , covPres=TRUE, CovarFile="covar.txt")
```

`Rlinear` *Robust SNP association tests for quantitative outcomes allowing for covariates*

Description

This function performs robust genetic association tests for quantitative outcomes. SNP associations are tested under additive, dominant and recessive models. Valid p-values are returned for each SNP, taking into account the multiple testing under different inheritance models. Covariates may also be included.

Usage

```
Rlinear(data = NA, outcome = NA, SNP = NA, InputFile = NA, covPres = TRUE,
COVAR = NA, CovarFile = NA, inputheader = TRUE, covarheader = TRUE,
maxpts = 25000, abseps = 0.001)
```

Arguments

`data`	A matrix or dataframe specifying the genotypes. Each row represents an individual and each column represents a SNP. SNPs should be coded as 0,1 or 2 according to the allelic counts.
`outcome`	The quantitative outcome.
`SNP`	SNP labels. If not specified, SNPs will be labelled by numbers.
`InputFile`	An optional txt file specifying the outcome in the first column and the genotypes in the other columns. Each row represents an individual. SNPs should be coded as 0,1 or 2 according to the allelic counts. Users may either specify both `data` and `outcome` or just specify `InputFile`.
`covPres`	Whether covariates are present. If `covPres` is `TRUE`, the user must specify either `COVAR` OR `CovarFile`.
`COVAR`	(Optional) A matrix or dataframe specifying the covariates. Each row represents an individual and each column represents a covariate.
`CovarFile`	(Optional) A file location for covariates. The file should be readable by the `read.table` command in R.
`inputheader`	(Optional) Whether `InputFile` contains headers. Only applies when `InputFile` is given.
`covarheader`	(Optional) Whether `CovarFile` contains headers. Only applies when `CovarFile` is given.
`maxpts`	(Optional) Arguments for `pmvnorm` in the `mvtnorm` package. Specifies the maximum number of function values (as integer).
`abseps`	(Optional) Arguments for `pmvnorm` in the `mvtnorm` package. Specifies the absolute error tolerance (as double).

Details

See `Rbin`.

Value

Same as the ouput from Rbin.

Author(s)

Hon-Cheong So

References

Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

See Also

Rlinear.block

Examples

```
## Simuate a case-control study with 300 cases/300 controls
## and 10 SNPs
set.seed(3)
outcome1 = rnorm(600,mean=30,sd=10)
data1 = matrix(nrow=600, ncol=10)
for (k in 1:10) {
data1[,k]=sample ( rep(0:2,200) ) }
COVAR1= rnorm(600,mean=100,sd=30)

## with covariates
Rlinear( data= data1, outcome=outcome1, SNP=1:10,covPres=TRUE,
COVAR= COVAR1)

##without covariates
Rlinear( data= data1, outcome=outcome1, SNP=1:10,covPres=FALSE)
```

Rlinear.block — *Robust SNP association tests for quantitative outcomes allowing for covariates, with genotypes read in blocks*

Description

This function is similar to Rlinear, but the genotypes are read in blocks (instead of reading all of them simultaneously). The aim to to reduce the memory consumption. This function is particularly suited for large-scale association studies like genome-wide association studies (GWAS).

Usage

```
Rlinear.block(GenoFile = NA, covPres = TRUE, COVAR = NA, CovarFile = NA,
SNPrange = NA, covarheader = TRUE, blocksize = 5000, maxpts = 25000,
abseps = 0.001)
```

Arguments

GenoFile	A file location for genotype file in PLINK binary PED format. A collection of three files with the same filename but different suffices(bed, bim and fam) should be present. Only the filename but NOT the suffix should be given.
covPres	Whether covariates are present. If covPres is TRUE, the user must specify either COVAR OR CovarFile.
COVAR	(Optional) A matrix or dataframe specifying the covariates. Each row represents an individual and each column represents a covariate.
CovarFile	(Optional) A file location for covariates. The file should be readable by the read.table command in R.
SNPrange	(Optional) A numeric vector specifying the range of SNPs to be tested.
covarheader	(Optional) Whether CovarFile contains headers. Only applies when CovarFile is given.
blocksize	The number of SNPs to be read at one time.
maxpts	(Optional) Arguments for pmvnorm in the mvtnorm package. Specifies the maximum number of function values (as integer).
abseps	(Optional) Arguments for pmvnorm in the mvtnorm package. Specifies the absolute error tolerance (as double).

Details

See Rbin.block .

Value

Same as the ouput from Rbin.

Author(s)

Hon-Cheong So

References

Hon-Cheong So and Pak C. Sham (2010).Robust association tests under different genetic models, allowing for binary or quantitative traits and covariates. Submitted.

See Also

Rlinear

Examples

```
## Assume the binary PED file is named "gwas" and the covariate file is "covar.txt"
## Result = Rlinear.block (GenoFile="gwas" , covPres=TRUE, CovarFile="covar.txt")
```

Index