

Package ‘ACTGE’

September 13, 2016

Title Adaptive Combination Testing and Multiple Testing for Gene-Environment Interactions

Version 1.0

Description Test Interaction between a Set of SNPs and an Exposure Factor

Depends R (>= 3.0.2), mvtnorm

License USC

LazyData true

RoxygenNote 5.0.1

R topics documented:

GE_ACT	1
GE_multiple	3
Index	5

GE_ACT	<i>Adaptive Combination Tests for Set-Based Gene-Environment Interactions</i>
--------	---

Description

This function returns P values of the interaction effect between a set of SNPs and exposure status. P values include ACT, ACT weighted by screening statistics in both Z scale and Chi-Square scale. Users need to input a set of genetic variants G, outcome Y, exposure E, and a few covariates to adjust W and outcome type, "B" for binary and "C" for continuous.

Usage

```
GE_ACT(G, Y, E, W, outcomeType, GeneID = NULL, maf_cut = 0,
       missing_cut = 0.15, nReps = 10000)
```

Arguments

G	n by p matrix of p SNPs for n subjects
Y	length n vector of outcome for n subjects
E	length n vector of exposure for n subjects

W	n by q data frame of q covariates to adjust for n subjects
outcomeType	"B" for binary outcome and "C" for continuous outcome
GeneID	length p vector indicating GeneID for each SNP, i.e which Gene does each SNP belong to? If you don't want a two-level analysis, you can ignore it
maf_cut	SNPs with Minor Allele Frequency lower than this threshold would be removed from the set, by default 0, i.e no SNP removed
missing_cut	SNPs with missing rate higher than this cut-off value would be removed from the set, by default 0.15. Missing values would be imputed by average
nReps	number of resampling to calculate ACT P Value, by default 10000. We recommend to use a small value, say 1000, to see the significance. If significant, then try larger value to get more accurate P Values.

Details

The input G, Y, E, W need to be numerical, with same number of observations. Note that Y and E cannot contain missing values, and please remove those observations with missing in Y or E. We suggest users impute G before using this function, though SNPs with missing rate lower than missing_cut would be imputed in the function with mean value for that SNP.

Value

nGene:	number of Genes
P.Values:	A list of ACT P values
p.ACT.Z.weight.pos	ACT P value based on positive part of marginal Z statistics weighted by screening statistics
p.ACT.Z	ACT P value based on marginal Z statistics weighted by screening statistics
p.ACT.Z.weight	ACT P value based on marginal Z statistics
p.ACT.X	ACT P value based on marginal Chi-Square statistics
p.ACT.X.weight	ACT P value based on marginal Chi-Square statistics weighted by screening statistics
Z.Statistics:	Z Statistics used to calculate ACT P Values
GE.interaction	Z statistics of marginal GE interaction
GY.association	Z statistics of marginal GY association
GE.correlation	Z statistics of marginal GE correlation
Weight	screening Z statistics used as weight, stronger of GY association/GE correlation
GE.interaction, GY.association, GE.correlation summary	
Estimate	Coefficient Estimate
Std.Error	Standard Error
z value	Z statistics
Pr(> z)	P Value

Author(s)

Shirong Zhang <shirongz@usc.edu>

Juan Pablo Lewinger <lewinger@usc.edu>

References

Shirong Zhang, Juan Pablo Lewinger Adaptive set-based methods for Gene-Environment Interactions

Examples

```
## Generate data for examples
set.seed(100)
G1=rbinom(2000,1,0.3)
G2=rbinom(2000,1,0.4)
G3=rbinom(2000,1,0.25)
G4=rbinom(2000,1,0.5)
G=cbind(G1,G2,G3,G4)
E=rbinom(2000,1,0.3)
Y=rbinom(2000,1,0.2)
GeneID=c("Gene1","Gene1","Gene2","Gene3")
## Example without covariates to adjust
GE_ACT(G,Y,E,W=NULL,outcomeType="B",GeneID)

## Example with covariates to adjust
age=rnorm(2000,45,2.6)
bmi=rnorm(2000,11.2,2.2)
GE_ACT(G,Y,E,W=cbind.data.frame(age,bmi),outcomeType="B")

## Example if users want to change default settings
Y=rnorm(2000,2,4)
GE_ACT(G,Y,E,W=cbind.data.frame(age,bmi),outcomeType="C",maf_cut=0.05,missing_cut=0.1)
```

GE_multiple

Multiple Testing for Gene-Environment Interactions (MTGE)

Description

This function returns P values of the interaction effect between individual SNPs and exposure status. P values include unadjusted P values, P values adjusted by MTGE and Sidak. Users need to input a set of genetic variants G, outcome Y, exposure E, and a few covariates to adjust W and outcome type, "B" for binary and "C" for continuous.

Usage

```
GE_multiple(G, Y, E, W, outcomeType, maf_cut = 0, missing_cut = 0.15)
```

Arguments

G	n by p matrix of p SNPs for n subjects
Y	length n vector of outcome for n subjects
E	length n vector of exposure for n subjects
W	n by q data frame of q covariates to adjust for n subjects
outcomeType	"B" for binary outcome and "C" for continuous outcome
maf_cut	SNPs with Minor Allele Frequency lower than this threshold would be removed from the set, by default 0, i.e no SNP removed
missing_cut	SNPs with missing rate higher than this cut-off value would be removed from the set, by default 0.15. Missing values would be imputed by average

Details

The input G, Y, E, W need to be numerical, with same number of observations. Note that Y and E cannot contain missing values, and please remove those observations with missing in Y or E. We suggest users impute G before using this function, though SNPs with missing rate lower than missing_cut would be imputed in the function with mean value for that SNP.

Value

unadjusted and adjusted P values for each individual SNP

p_unadjsted	unadjusted P values
p_MTGE	adjusted P values using MTGE
p_Sidak	adjusted P values using Sidak

Author(s)

Shirong Zhang <shirongz@usc.edu>

Juan Pablo Lewinger <lewinger@usc.edu>

References

Shirong Zhang, Juan Pablo Lewinger Multiple Testing for Gene-Environment Interactions

Examples

```
## Generate data for examples
set.seed(100)
G1=rbinom(2000,1,0.3)
G2=rbinom(2000,1,0.4)
G3=rbinom(2000,1,0.25)
G4=rbinom(2000,1,0.5)
G=cbind(G1,G2,G3,G4)
E=rbinom(2000,1,0.3)
Y=rbinom(2000,1,0.2)

## Example without covariates to adjust
GE_multiple(G,Y,E,W=NULL,outcomeType="B")

## Example with covariates to adjust
age=rnorm(2000,45,2.6)
bmi=rnorm(2000,11.2,2.2)
GE_multiple(G,Y,E,W=cbind.data.frame(age,bmi),outcomeType="B")

## Example if users want to change default settings
Y=rnorm(2000,2,4)
GE_multiple(G,Y,E,W=cbind.data.frame(age,bmi),outcomeType="C",maf_cut=0.05,missing_cut=0.1)
```

Index

GE_ACT, [1](#)
GE_multiple, [3](#)