

Package ‘MPAT’

October 5, 2015

Type Package

Title This package is for multiple phenotype genetic association studies

Version 1.0

Date 2015-09-10

Author Zhonghua Liu

Maintainer Zhonghua Liu <zliu@mail.harvard.edu>

Description Conduct multiple phenotype genetic association studies based on summary statistics.

License GPL (≥ 2)

NeedsCompilation yes

Imports mvtnorm

Archs i386, x64

R topics documented:

MPAT-package	2
DSUM	2
lipids	3
MinP	4
mixAda	5
mixFisher	6
mixSDWeight	7
mixTippett	7
mixVarWeight	8
PC	9
PCFisher	10
PCLC	11
PCLogit	11
PCMinP	12
SUM	13
VC	14
Wald	14
WI	15
Index	17

MPAT-package

Multiple Phenotype Genetic Association Testings by Combining Univariate Testing Summary Statistics.

Description

This package is for conducting multiple phenotype genetic association testings by combining univariate summary statistics for each phenotype in GWAS. It contains fourteen methods that can be used in a variety of situations. An overall p-value will be returned for assessing the significance of associations between a SNP and multiple phenotypes. It contains the following functions: PC, PCMinP, PCFisher, PCLogit, PCLC, PCSS, MinP, WI, VC, DSUM, SUM, Wald, mixAda, mixFisher, mixTippett, mixVarWeight, mixSDWeight.

Details

Package: MPAT
Type: Package
Version: 1.0
Date: 2014-09-11
License: License: GPL(>= 2)

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>
Maintainer: Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

DSUM

Direct summation of Z-statistics

Description

For SNP, the Z testing statistics for multiple phenotypes are summed together as an overall testing statistic.

Usage

```
DSUM(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics that can be estimated by its sample version over the SNPs.

Value

p-value An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
DSUM(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

lipids	<i>An example data set that contains summary statistics from GWAS studies of four lipids levels.</i>
--------	--

Description

This is an example data set that contains summary statistics (Z-scores) for 2000 genetic variants and four lipids levels.

Usage

```
data(lipids)
```

Format

A data frame with 2000 observations on the following 15 variables.

MarkerName SNP names

Allele1 This the allele used as the effect allele

Allele2 This is the "other" allele

W.HDL The sum of the individual study weights (typically, N) for this marker

Zscore.HDL Z scores for HDL

Pvalue.HDL P-values for HDL

W.LDL The sum of the individual study weights (typically, N) for this marker

Zscore.LDL Z scores for LDL

Pvalue.LDL P-values for LDL

W.TC The sum of the individual study weights (typically, N) for this marker

Zscore.TC Z scores for TC

Pvalue.TC P-values for TC

W.TG The sum of the individual study weights (typically, N) for this marker

Zscore.TG Z scores for TG

Pvalue.TG P-values for TG

Source

<http://www.sph.umich.edu/csg/abecasis/public/lipids2010/>

References

<http://www.nature.com/nature/journal/v466/n7307/full/nature09270.html>

Examples

```
data(lipids)
head(lipids)
```

MinP

MinP takes the minimum of the p-values for multiple phenotypes as the overall testing statistic.

Description

MinP uses the minimum of the p-values for multiple phenotypes as the overall testing statistic. The p-value of MinP is calculated based on multivariate normal distribution.

Usage

```
MinP(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
MinP(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

mixAda

Adaptively select optimal combination coefficient of two independent-score statistics based on linear mixed effects model.

Description

mixAda combines the Z testing statistics for multiple phenotypes at a genetic variant in an adaptive fashion to maximize the power.

Usage

```
mixAda(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
mixAda(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

mixFisher	<i>Combine Z-statistics for multiple phenotypes using Fisher's method in a linear mixed effects model.</i>
-----------	--

Description

Z-statistics for multiple phenotypes are modeled using a linear mixed effects model. Two independent score testing statistics are combined using Fisher's method.

Usage

```
mixFisher(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p_group	p-value for the group effects of a genetic variant on the multiple phenotypes
p_individual	p-value for the individual effects of a genetic variant on the multiple phenotypes
p_overall	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
mixFisher(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

mixSDWeight	<i>Combine Z-statistics for multiple phenotypes using inverse standard deviation weighting method in a linear mixed effects model.</i>
-------------	--

Description

Z-statistics for multiple phenotypes are modeled using a linear mixed effects model. Two independent score testing statistics are combined using inverse standard deviation weighting method.

Usage

```
mixSDWeight(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
mixSDWeight(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

mixTippett	<i>Combine Z-statistics for multiple phenotypes using Tippett's method in a linear mixed effects model.</i>
------------	---

Description

Z-statistics for multiple phenotypes are modeled using a linear mixed effects model. Two independent score testing statistics are combined using Tippett's method.

Usage

```
mixTippett(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p_group	p-value for the group effects of a genetic variant on the multiple phenotypes
p_individual	p-value for the individual effects of a genetic variant on the multiple phenotypes
p_overall	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
mixTippett(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

mixVarWeight	<i>Combine Z-statistics for multiple phenotypes using inverse variance weighting method in a linear mixed effects model.</i>
--------------	--

Description

Z-statistics for multiple phenotypes are modeled using a linear mixed effects model. Two independent score testing statistics are combined using inverse variance weighting method.

Usage

```
mixVarWeight(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
mixVarWeight(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

PC

PC is a function used to combine correlated GWAS summary statistics for multiple phenotypes at a genetic locus.

Description

By combining univariate summary statistics for multiple phenotypes, we can obtain an overall association p-value that assess the association between a genetic variant and multiple phenotypes as a whole. PC allows users to conduct dimension reduction by specifying which principal component to use to represent the original multiple summary statistics.

Usage

```
PC(Z.vec, Sigma, PCorder)
```

Arguments

Z.vec	Z.vec is a column vector representing the Wald-type Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix of the multiple Z testing statistics.
PCorder	PCorder specifies which principal component to use, it takes integer values. For example, PCorder=1 means we choose to use the first principal component.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
PC(Z.vec=lipids_zscore[1,],Sigma,PCorder=1) ## p-value for the first SNP using PC1
```

PCFisher

Fisher's method to combine principal components based on Z testing statistics

Description

Z testing statistics for multiple phenotype are orthogonally transformed onto principal component axes and combined using Fisher's method.

Usage

```
PCFisher(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
PCFisher(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

PCLC	<i>Linear combination of the principal components of the Z testing statistics.</i>
------	--

Description

Z testing statistics for multiple phenotype are orthogonally transformed onto principal component axes and combined in a linear fashion.

Usage

```
PCLC(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
PCLC(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

PCLogit	<i>Logit method to combine principal components based on Z testing statistics</i>
---------	---

Description

Z testing statistics for multiple phenotype are orthogonally transformed onto principal component axes and combined using Logit method.

Usage

```
PCLogit(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
PCLogit(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

PCMinP

Use the minimum p-value of the principal components of the Z testing statistics as the testing statistic

Description

The p-values using principal components of the Z testing statistics are computed, and the minimum one is taken as the testing statistic.

Usage

```
PCMinP(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
PCMinP(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

SUM	<i>Use SUM method to combine Z testing statistics</i>
-----	---

Description

The Z testing statistics for multiple phenotypes are combined using SUM method with correlation structures explicitly taken into account.

Usage

```
SUM(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
SUM(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

VC	<i>Use variance component test (VC) to combine Z testing statistics</i>
----	---

Description

Use variance component test (VC) to combine Z testing statistics of multiple phenotypes.

Usage

```
VC(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
VC(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

Wald	<i>Traditional Wald type combination of Z testing statistics</i>
------	--

Description

The Z testing statistics for multiple phenotypes are combined using Wald type method.

Usage

```
Wald(Z.vec, Sigma)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
Wald(Z.vec=lipids_zscore[1,],Sigma) ## p-value for the first SNP
```

WI

Use WI method to combine Z testing statistics.

Description

Use WI method to combine Z testing statistics.

Usage

```
WI(Z.vec, Sigma,method)
```

Arguments

Z.vec	Z.vec is a vector representing the Z testing statistics for multiple phenotypes.
Sigma	Sigma is the correlation matrix among the Z testing statistics.
method	specify which method to be used to compute the p-value

Value

p-value	An overall p-value assessing the association between a genetic variant and multiple phenotypes is returned.
---------	---

Author(s)

Zhonghua Liu <zliu@mail.harvard.edu>

References

MPAT: an R package for multivariate phenotype association studies.

Examples

```
data(lipids)
lipids = as.data.frame(lipids)
lipids_zscore = as.matrix(lipids[,c("Zscore.HDL", "Zscore.LDL", "Zscore.TG", "Zscore.TC")])
Sigma = cor(lipids_zscore )
WI(Z.vec=lipids_zscore[1,],Sigma,method="liu") ## p-value for the first SNP
```


Index

*Topic **datasets**

lipids, [3](#)

DSUM, [2](#)

lipids, [3](#)

MinP, [4](#)

mixAda, [5](#)

mixFisher, [6](#)

mixSDWeight, [7](#)

mixTippett, [7](#)

mixVarWeight, [8](#)

MPAT (MPAT-package), [2](#)

MPAT-package, [2](#)

PC, [9](#)

PCFisher, [10](#)

PCLC, [11](#)

PCLogit, [11](#)

PCMinP, [12](#)

PCSS (Wald), [14](#)

SUM, [13](#)

VC, [14](#)

Wald, [14](#)

WI, [15](#)