Package 'MiST'

February 19, 2015

Type Package

Title Mixed effects Score Test for continuous outcomes

Version 1.0

Date 2013-03-18

Author Jianping Sun, Yingye Zheng, and Li Hsu

Maintainer Jianping Sun <jsun@fhcrc.org>

Depends CompQuadForm

Description Test for association between a set of SNPS/genes and continuous or binary outcomes by including variant characteristic information and using (weighted) score statistics.

License LGPL (>= 2.0)

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2013-12-14 17:01:19

R topics documented:

MiST-package *Mixed effects Score Test*

Description

Test for association between a set of SNPS/genes and continuous or binary outcomes by including variant characteristic information and using score statistics.

Details

Package:	MiST
Type:	Package
Version:	1.0
Date:	2013-03-18
License:	LGPL (>= 2.0)
LazyLoad:	yes

linear.test(y, X, G, Z, method = "liu")

linear.weight.test(y, X, G, Z, maf, weight.beta = c(1, 25), method = "liu")

logit.test(y, X, G, Z, method = "liu")

logit.weight.test(y, X, G, Z, maf, weight.beta = c(1, 25), method = "liu")

Author(s)

Jianping Sun, Yingye Zheng, and Li Hsu.

References

Sun, J., Zheng, Y., and Hsu, L. (2013) A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies. Genet Epidemiol. 2013 Mar 9. doi: 10.1002/gepi.21717

H. Liu, Y. Tang, H.H. Zhang (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, Computational Statistics and Data Analysis, 53, 853-856.

linear.test *Mixed effects Score Test for continuous outcomes*

Description

Test for association between a set of SNPS/genes and continuous outcomes by including variant characteristic information and using score statistics.

Usage

```
linear.test(y, X, G, Z, method = "liu")
```

Arguments

y	a numeric vector of the continuous outcome variables. Missing values are not allowed.
X	a numeric matrix of covariates with rows for individuals and columns for covariates.
G	a numeric genotype matrix with rows for individuals and columns for SNPs. Each SNP should be coded as 0, 1, and 2 for AA, Aa, aa, where A is a major allele and a is a minor allele. Missing genotypes are not allowed.
Z	a numeric matrix of second level covariates for variant characteristics. Each row corresponds to a variant and each column corresponds to a variant characteristic. If there is no second level covariates, a vector of 1 should be used.
method	a method to compute the p-value and the default value is "liu". Method "davies" represents an exact method that computes the p-value by inverting the characteristic function of the mixture chisq. Method "liu" represents an approximation method that matches the first 3 moments.

Value

S.tau	score statistic for the variant hetergenous effect.
S.pi	score statistic for the variant mean effect.
p.value.S.tau	p-value for testing the variant hetergenous effect.
p.value.S.pi	p-value for testing the variant mean effect.
p.value.overall	overall p-value for testing the association between the set of SNPS/genes and outcomes. It combines p.value.S.pi and p.value.S.tau by using Fisher's procedure.

Author(s)

Jianping Sun, Yingye Zheng, and Li Hsu.

References

Sun, J., Zheng, Y., and Hsu, L. (2013) A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies. Genet Epidemiol. 2013 Mar 9. doi: 10.1002/gepi.21717

H. Liu, Y. Tang, H.H. Zhang (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, Computational Statistics and Data Analysis, 53, 853-856.

Examples

```
data(MiST.data)
attach(MiST.data)

#########################################################################
# Test the association between a set of SNPs and continuous outcomes
#   - without information about SNP characteristics. Z is a vector of 1's.


out <- linear.test(y.con, X, G, Z)


#########################################################################
# Test the association between a set of SNPs and continuous outcomes
# - including SNP characteristics


out <- linear.test(y.con, X, G, Z.func)
```

`linear.weight.test` *Weighted Mixed effects Score Test for continuous outcomes*

Description

Test for association between a set of SNPS/genes and continuous outcomes by including variant characteristic information and using weighted score statistics.

Usage

```
linear.weight.test(y, X, G, Z, maf, weight.beta = c(1, 25), method = "liu")
```

Arguments

`y`	a numeric vector of the continuous outcome variables. Missing values are not allowed.
`X`	a numeric matrix of covariates with rows for individuals and columns for covariates.
`G`	a numeric genotype matrix with rows for individuals and columns for SNPs. Each SNP should be coded as 0, 1, and 2 for AA, Aa, aa, where A is a major allele and a is a minor allele. Missing genotypes are not allowed.
`Z`	a numeric matrix of second level covariates for variant characteristics. Each row corresponds to a variant and each column corresponds to a variant characteristic. If there is no second level covariates, a vector of 1 should be used.
`maf`	a numeric vector of MAF (minor allele frequency) for each SNP.

`weight.beta` a numeric vector of parameters of beta function which is the weight for score statistics. The default value is "c(1,25)".

`method` a method to compute the p-value and the default value is "liu". Method "davies" represents an exact method that computes the p-value by inverting the characteristic function of the mixture chisq. Method "liu" represents an approximation method that matches the first 3 moments.

Value

`S.tau` score statistic for the variant hetergenous effect.

`S.pi` score statistic for the variant mean effect.

`p.value.S.tau` p-value for testing the variant hetergenous effect.

`p.value.S.pi` p-value for testing the variant mean effect.

`p.value.overall` overall p-value for testing the association between the set of SNPS/genes and outcomes. It combines p.value.S.pi and p.value.S.tau by using Fisher's procedure.

Author(s)

Jianping Sun, Yingye Zheng, and Li Hsu.

References

Sun, J., Zheng, Y., and Hsu, L. (2013) A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies. Genet Epidemiol. 2013 Mar 9. doi: 10.1002/gepi.21717

H. Liu, Y. Tang, H.H. Zhang (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, Computational Statistics and Data Analysis, 53, 853-856.

Examples

```
data(MiST.data)
attach(MiST.data)

##########################################################################
# Test the association between a set of SNPs and continuous outcomes
#   - without information about SNP characteristics. Z is a vector of 1's.


out <- linear.weight.test(y.con, X, G, Z, maf)


##########################################################################
# Test the association between a set of SNPs and continuous outcomes
# - including SNP characteristics


out <- linear.weight.test(y.con, X, G, Z.func, maf)
```

`logit.test` *Mixed effects Score Test for binary outcomes*

Description

Test for association between a set of SNPS/genes and binary outcomes by including variant characteristic information and using score statistics.

Usage

```
logit.test(y, X, G, Z, method = "liu")
```

Arguments

`y`	a numeric (0 or 1) vector of the binary outcome variables. Missing values are not allowed.
`X`	a numeric matrix of covariates with rows for individuals and columns for covariates.
`G`	a numeric genotype matrix with rows for individuals and columns for SNPs. Each SNP should be coded as 0, 1, and 2 for AA, Aa, aa, where A is a major allele and a is a minor allele. Missing genotypes are not allowed.
`Z`	a numeric matrix of second level covariates for variant characteristics. Each row corresponds to a variant and each column corresponds to a variant characteristic. If there is no second level covariates, a vector of 1 should be used.
`method`	a method to compute the p-value and the default value is "liu". Method "davies" represents an exact method that computes the p-value by inverting the characteristic function of the mixture chisq. Method "liu" represents an approximation method that matches the first 3 moments.

Value

`S.tau`	score statistic for the variant hetergenous effect.
`S.pi`	score statistic for the variant mean effect.
`p.value.S.tau`	p-value for testing the variant hetergenous effect.
`p.value.S.pi`	p-value for testing the variant mean effect.
`p.value.overall`	overall p-value for testing the association between the set of SNPS/genes and outcomes. It combines p.value.S.pi and p.value.S.tau by using Fisher's procedure.

Author(s)

Jianping Sun, Yingye Zheng, and Li Hsu.

References

Sun, J., Zheng, Y., and Hsu, L. (2013) A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies. Genet Epidemiol. 2013 Mar 9. doi: 10.1002/gepi.21717

H. Liu, Y. Tang, H.H. Zhang (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, Computational Statistics and Data Analysis, 53, 853-856.

Examples

```
data(MiST.data)
attach(MiST.data)

######################################################################
# Test the association between a set of SNPs and binary outcomes
#   - without information about SNP characteristics. Z is a vector of 1's.

out <- logit.test(y.bin, X, G, Z)


######################################################################
# Test the association between a set of SNPs and bianry outcomes
# - including SNP characteristics

out <- logit.test(y.bin, X, G, Z.func)
```

`logit.weight.test` *Weighted Mixed effects Score Test for binary outcomes*

Description

Test for association between a set of SNPS/genes and binary outcomes by including variant characteristic information and using weighted score statistics.

Usage

```
logit.weight.test(y, X, G, Z, maf, weight.beta = c(1, 25), method = "liu")
```

Arguments

`y`	a numeric vector (0 or 1) of the binary outcome variables. Missing values are not allowed.
`X`	a numeric matrix of covariates with rows for individuals and columns for covariates.
`G`	a numeric genotype matrix with rows for individuals and columns for SNPs. Each SNP should be coded as 0, 1, and 2 for AA, Aa, aa, where A is a major allele and a is a minor allele. Missing genotypes are not allowed.

Z	a numeric matrix of second level covariates for variant characteristics. Each row corresponds to a variant and each column corresponds to a variant characteristic. If there is no second level covariates, a vector of 1 should be used.
maf	a numeric vector of MAF (minor allele frequency) for each SNP.
weight.beta	a numeric vector of parameters of beta function which is the weight for score statistics. The default value is "c(1,25)".
method	a method to compute the p-value and the default value is "liu". Method "davies" represents an exact method that computes the p-value by inverting the characteristic function of the mixture chisq. Method "liu" represents an approximation method that matches the first 3 moments.

Value

S.tau	score statistic for the variant hetergenous effect.
S.pi	score statistic for the variant mean effect.
p.value.S.tau	p-value for testing the variant hetergenous effect.
p.value.S.pi	p-value for testing the variant mean effect.
p.value.overall	overall p-value for testing the association between the set of SNPS/genes and outcomes. It combines p.value.S.pi and p.value.S.tau by using Fisher's procedure.

Author(s)

Jianping Sun, Yingye Zheng, and Li Hsu.

References

Sun, J., Zheng, Y., and Hsu, L. (2013) A Unified Mixed-Effects Model for Rare-Variant Association in Sequencing Studies. Genet Epidemiol. 2013 Mar 9. doi: 10.1002/gepi.21717

H. Liu, Y. Tang, H.H. Zhang (2009) A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, Computational Statistics and Data Analysis, 53, 853-856.

Examples

```
data(MiST.data)
attach(MiST.data)

#########################################################################
# Test the association between a set of SNPs and binary outcomes
#   - without information about SNP characteristics. Z is a vector of 1's.

out <- logit.weight.test(y.bin, X, G, Z, maf)


#########################################################################
# Test the association between a set of SNPs and bianry outcomes
```

```
# - including SNP characteristics

out <- logit.weight.test(y.bin, X, G, Z.func, maf)
```

MiST.data *Data Example for MiST*

Description

A numerical data example for MiST

Usage

```
data(MiST.data)
```

Format

MiST.data contains the following objects:

- G: a numeric genotype matrix of 3400 individuals and 93 SNPs. Each row represents a individual, and each column represents a SNP marker.
- X: a numeric matrix of covariates with the first column represents intercept, the second column represents a continuous covariate, and the third column represents a binary covariate.
- Z: a numeric vector of 1s for the second level covariate.
- Z.func: a numeric matrix for the second level covariate. The first column contains all 1s, representing the intercept, and the second column is 0 or 1, representing whether a SNP marker is functional or non-functional.
- maf: a numeric vector for the Minor Allele Frequecy of 93 SNPs.
- y.con: a numeric vector of continuous outcomes.
- y.bin: a numeric vector of binary outcomes.

Examples

```
data(MiST.data)
```

Index