Package ‘gskat’

February 20, 2015

Type Package
Title GEE_KM
Version 1.0
Date 2012-12-20
Author Xuefeng Wang
Maintainer Xuefeng Wang <pwxf@gmail.com>
Description Family based association test via GEE Kernel Machine score test
License GPL (>= 2)
LazyData yes
LazyLoad yes
Depends R (>= 2.10), CompQuadForm, e1071, gee, geepack, Matrix
NeedsCompilation no
Repository CRAN
Date/Publication 2013-11-04 08:07:19

R topics documented:

  gskat-package .......................................................... 2
gdata ................................................................. 2
gskat_score ........................................................... 3
gskat_score_pert ....................................................... 3
gskat_seq ............................................................... 4
gskat_seq_cont ......................................................... 5
score.RFAM_IC_burden ............................................... 6
score_FSKAT_IC_pertu ............................................... 7

Index 9
**gskat-package**

*Family based association test via kernel machine regression score test*

### Description

R package for GEE_SKAT.

### Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>gskat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.0</td>
</tr>
<tr>
<td>Date:</td>
<td>2012-12-30</td>
</tr>
<tr>
<td>License:</td>
<td>GPL-2</td>
</tr>
<tr>
<td>LazyLoad:</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Author(s)

XUEFENG WANG  
Maintainer: Xuefeng Wang <pxwx@gmail.com>

---

**gdata**

*Sample Data*

### Description

Sample data

### Usage

data(gdata)

### Format

The format is:
Perform GEE_SKAT Score Test

Description

This function performs family based marker set based association (score test) based on the GEE_SKAT.

Usage

```r
gskat_score(peddat,F1=FALSE)
```

Arguments

- **pedDat**: Input Family Data: A list including four data matrix: ID, y, X, Z. ID: Pedigree ID matrix, having four columns are the Family ID (FID), Individual ID (IID), Paternal ID (FAT), Maternal ID (MAT), respectively. y: binary phenotype coded as 0, 1. X: covariate matrix, including the intercept column Z: SNP genotypes coded 0/1/2 (minor allele count).
- **F1**: If TRUE, use identity working corr. matrix

Value

- **pval**: Returns p-value
- **ifault**: fault indicator from the davies method: 0: no error, 1: requested accuracy could not be obtained, 2: round-off error possibly significant, 3: invalid parameters, 4: unable to locate integration parameters

Author(s)

Xuefeng Wang

Examples

```r
gskat_score(gdata)
```

Perform GEE_SKAT Score Test (perturbation)

Description

This function performs family based marker set based association (score test) based on the GEE_SKAT.

Usage

```r
gskat_score_pert(pedDat,F1=FALSE,pw="Rade",np=10000)
```
Arguments

pedDat: Input Family Data: A list including four data matrix: ID, y, X, Z. ID: Pedigree ID matrix. The four columns are the Family ID (FID), Individual ID (IID), Paternal ID (FAT), Maternal ID (MAT), respectively. y: binary phenotype coded as 0, 1 X: covariate matrix, including the intercept column Z: SNP genotypes coded 0/1/2 (minor allele count).

F1: If TRUE, use identity working corr. matrix

pw: Perturbation method. "Rade": Rademacher; "Norm": Normal distribution

np: number of perturbed samples. default=10000

Value

pval: Returns p-value
Ts: Score statistic
mu_Ts: Theoretical mean of Ts
var_Ts: Theoretical variance of Ts
PM: Empirical mean of Ts based on the perturbation
PV: Empirical variance of Ts based on the perturbation

Author(s)

Xuefeng Wang

Examples

gskat_score_pert(gdata)

Description

This function performs family based marker set based association (score test) based on the GEE_KM.

Usage

gskat_seq(y, X, Z, ID, impute.method="fixed", SNP.weights=NULL, w_a=1, w_b=25, resampling=TRUE, pw="Rade", Uc=TRUE, sW=FALSE, np=10000)
**Arguments**

- **y**: binary phenotype coded as 0, 1
- **XC**: covariate matrix, including the intercept column
- **Z**: SNP genotypes coded 0/1/2 (minor allele count)
- **ID**: ID: Pedigree ID matrix, having four columns are the Family ID (FID), Individual ID (IID), Paternal ID (FAT), Maternal ID (MAT), respectively.
- **impute.method**: default is fixed method i.e. fill with means
- **SNP.weights**: If NULL, the default beta (1,25) desensity will be used, or a custimoized weight vector may be used
- **w_a**: The first parameter of the beta density in the weight function
- **w_b**: The second parameter of the beta density in the weight function
- **resampling**: If TRUE, resampling will be applied
- **pw**: r.v. used in the perturbation, "Norm"=Normal , "Rade"=Rademacher
- **uc**: Score centered or not
- **sw**: standardize weights
- **np**: No. of perturbed samples

**Value**

- **p1**: Returns asymptotic p-value
- **p2**: Returns resampling p-value

**Author(s)**

Xuefeng Wang

**Examples**

```r
# gskat_score(gdata)$p2
```

---

gskat_seq_cont  
*Perform GEE_KM Score Test with sequencing data for continuous traits*

**Description**

This function performs family based marker set based association (score test) based on the GEE_KM.

**Usage**

```r
gskat_seq_cont(y, XC, Z, ID, impute.method="fixed", SNP.weights=NULL, w_a=1, w_b=25, resampling=TRUE, pw="Rade", uc=TRUE, sw=FALSE, np=10000)
```
Arguments

- **y**: binary phenotype coded as 0, 1
- **XC**: covariate matrix, including the intercept column
- **Z**: SNP genotypes coded 0/1/2 (minor allele count)
- **ID**: ID: Pedigree ID matrix, having four columns are the Family ID (FID), Individual ID (IID), Paternal ID (FAT), Maternal ID (MAT), respectively.
- **impute.method**: default is fixed method i.e. fill with means
- **SNP.weights**: If NULL, the default beta (1,25) density will be used, or a customized weight vector may be used
- **w_a**: The first parameter of the beta density in the weight function
- **w_b**: The second parameter of the beta density in the weight function
- **resampling**: If TRUE, resampling will be applied
- **pw**: r.v. used in the perturbation, "Norm"=Normal,"Rade"=Rademacher
- **uc**: Score centered or not
- **sw**: standardize weights or not
- **np**: No. of perturbed samples

Value

- **p1**: Returns asymptotic p-value
- **p2**: Returns resampling p-value

Author(s)

Xuefeng Wang

Examples

```r
# gskat_seq_cont(y, XC=X, Z, ID)
```

---

**score.RFAM_IC_burden**  
**Burden test based on GEE_SKAT**

Description

This function performs burden test based on the GEE_SKAT.

Usage

```r
score.RFAM_IC_burden(y, XC, Z, FID)
```
score_FSKAT_IC_pertu

Arguments

  y        binary phenotype coded as 0, 1
  xC       covariate matrix, WITHOUT the intercept column.
  Z        SNP genotypes coded 0/1/2 (minor allele count).
  FID      Family ID

Value

  Returns p-value

Author(s)

Xuefeng Wang

See Also

  geeglm_wrap; gee_wrap

Examples

  Z1=genoC(gdata$Z)
  score.RFAM.IC.burden(y=gdata$y, xC=gdata$x[, -1], Z=Z1, FID=gdata$ID[, 1])

Description

  GEE_SKAT for rare variants

Usage

  score_FSKAT_IC_pertu(y, xC, Z, ID, w_a=1, w_b=5, pw="Rade", uc=FALSE, sW=TRUE, np=10000)

Arguments

  y        binary phenotype coded as 0, 1
  xC       covariate matrix, WITHOUT the intercept column.
  Z        SNP genotypes coded 0/1/2 (minor allele count).
  ID       Pedigree ID matrix. The four columns are the Family ID (FID), Individual ID (IID), Paternal ID (FAT), Maternal ID (MAT), respectively.
  w_a      The first parameter of the beta density in the weight function
  w_b      The second parameter of the beta density in the weight function
  pw       perturbation method: "Rade"=Rademacher; "Norm"=Normal
  uc       Score centered or not
  sW       standardize weights or not
  np       No. of perturbed samples
Value

- `pval_davies`: Returns p-value from davies method
- `ifault`: fault indicator from the davies method: 0: no error, 1: requested accuracy could not be obtained, 2: round-off error possibly significant, 3: invalid parameters, 4: unable to locate integration parameters
- `PM`: Empirical mean of Ts based on perturbation
- `mu_Ts`: Theoretical mean of Ts
- `PV`: Empirical variance of Ts based on perturbation
- `var_Ts`: Theoretical variance of Ts
- `pval_pert`: Returns p-value from perturbation method

Author(s)

Xuefeng Wang

See Also

- `score.RFAM_IC_burden`

Examples

```r
score_FSKAT_IC_pertu(y=gdata$y, X=gdata$X[, -1], Z=gdata$Z, ID=gdata$ID)
```
Index

*Topic **datasets**
  gdata, 2
*Topic **package**
  gskat-package, 2

gdata, 2
 gskat (gskat-package), 2
 gskat-package, 2
 gskat_score, 3
 gskat_score_pert, 3
 gskat_seq, 4
 gskat_seq_cont, 5

 score.RFAM_IC_burden, 6
 score_FSKAT_IC_pertu, 7