

Package ‘LSRAT’

March 5, 2020

Type Package

Title Longitudinal variant-Set Retrospective Association Test

Version 0.0.0.9000

Date 2020-03-04

Author Weimiao Wu [aut, cre], Zuoheng Wang [aut]

Maintainer Weimiao Wu <weimiao.wu@yale.edu>

Description Perform set-based association test for longitudinal traits in genome-wide association studies for genetic rare variants.

Depends R (>= 3.5.1), snpStats, LBRAT, GENESIS, plyr, geepack, lme4, nlme, mvtnorm, knitr

License GPL (>= 2)

Encoding UTF-8

LazyData true

Date/Publication 2019-03-01

VignetteBuilder knitr

URL <https://github.com/WeimiaoWu/LSRAT>

BugReports <https://github.com/WeimiaoWu/LSRAT/issues>

RoxygenNote 7.0.2.9000

NeedsCompilation no

R topics documented:

lsrat.SSD.All	2
lsrat_est	2
lsrat_simu.bi	3
lsrat_simu.conti	3
lsrat_test	4
rsmmat_est	5
Index	6

lsrat.SSD.All	<i>LSRAT test or RSMMAT test using SSD format files</i>
---------------	---

Description

LSRAT test or RSMMAT test using SSD format files

Usage

```
lsrat.SSD.All(SSD.INFO, lsrat_est.obj, ...)
```

Arguments

SSD.INFO	SSD format information file, output of function "Open_SSD". The genome wide scan are run set by set.
lsrat_est.obj	ouput from lsrat_est() or rsmmat_est()
...	Other options of the LSRAT or RSMMAT test. Defined as same as in function "lsrat_test()" or "rsmmat_est()"

Value

results of the LSRAT test or RSMMAT test. First column contains batchID, second column contains SNP ID, third column contains prospective P-value and forth column contains retrospective P-value

lsrat_est	<i>GEE NULL model estimation</i>
-----------	----------------------------------

Description

This function estimate the parameters and residuals for the NULL model for GEE/LSRAT tests

Usage

```
lsrat_est(y.long, time, y.cov, timecov = TRUE, corstr = "ar1")
```

Arguments

y.long	Long-formatted phenotype vector
time	Time covarites matched with phenotype vector
y.cov	Covariate matrix denoting the covariate variables measured at each time
timecov	Logical variable, indicating whether the time fixed effect is estimated
corstr	String, correlation structure for GEE model, optional values are: 'ar1', 'ind', 'mixture'

Value

This function returns a list object with model parameters and residuals of the NULL GEE model

lsrat_simu.bi

*Simulation for LSRAT and RSMMAT test***Description**

This function use pre-defined parameters to make the simulation data with longitudinal binary traits for the LSRAT test (including type I and power test)

Usage

```
lsrat_simu.bi(
  n.sample = 5000,
  n.time = 7,
  par = list(),
  time_cov = TRUE,
  snp.count = 100,
  intercept = TRUE,
  power = FALSE
)
```

Arguments

n.sample	Numeric, sample size, number of individuals
n.time	Numeric, number of measurements for each individual
par	List, the parameters for the phenotype traits, including covaraites and individual specific time dependent random effects
time_cov	Logical variable, indicating whether time effect is included in phenotypic traits
snp.count	Numeric, number of SNPs in each variant set
intercept	Logical variable, indicating whether intercept is used in phenotypic traits
power	Logical variable, indicating whether the phenotype generated under the null model (type I error) or under the alternative model (power test)

Value

A list object is returned to be used as object for association test

lsrat_simu.conti

*Simulation for LSRAT and RSMMAT test***Description**

This function use pre-defined parameters to make the simulation data with longitudinal continuous traits for the LSRAT test (including type I and power test)

Usage

```
lsrat_simu.conti(
  n.sample = 5000,
  n.time = 7,
  par = list(),
  time_cov = TRUE,
  snp.count = 100,
  intercept = TRUE,
  power = FALSE,
  seed = NULL
)
```

Arguments

n.sample	Numeric, sample size, number of individuals
n.time	Numeric, number of measurements for each individual
par	List, the parameters for the phenotype traits, including covaraites and individual specific time dependent random effects
time_cov	Logical variable, indicating whether time effect is included in phenotypic traits
snp.count	Numeric, number of SNPs in each variant set
intercept	Logical variable, indicating whether intercept is used in phenotypic traits
power	Logical variable, indicating whether the phenotype generated under the null model (type I error) or under the alternative model (power test)

Value

A list object is returned to be used as object for association test

lsrat_test	<i>Calculate prospective and retrospective P-values for GEE or GLMM model</i>
------------	---

Description

This function tests a SNPs for a given SNP set for a given lbrat estimated null model.

Usage

```
lsrat_test(
  srat.est,
  G,
  weights = "beta",
  impute.method = "mean",
  GRM = NULL,
  tests = c("B", "S", "O", "E", "V", "R"),
  B = 5000,
  rho = c(0, 0.5, 1),
  return_single = FALSE
)
```

Arguments

srat.est	The output of function "lsrat_est()" or "rsmmat_est()"
G	The genotype matrix, an m*q matrix where m is the number of subjects and q is the total number genetic variants.
impute.method	choose the iputation method when there is missing genotype. Optional options are: 'random', 'fixed' or 'bestguess'.
GRM	takes m-by-m genetic correlation matrix or kinship matrix.
tests	a character vector indicating which LSRAT/RSMMAT tests should be performed ("B" for the burden test, "S" for SKAT, "C" for variant-level ACAT test, "O" for SKAT-O, "E" for the SMMAT test, "A" for omnibus ACAT test).
B	a number of perturbation used for P-value approximation
rho	a numeric vector defining the search grid used in SMMAT-O for SKAT-O (see the SKAT-O paper for details). Default = c(0, 0.5, 1)
return_single	Logical parameter indicating whether single variant P-value to be returned or not. Default = FALSE.

Value

This function returns a dataframe. The row name is the SNP ID, the first column is the prospective score statistics, the second column is the retrospective score statistics, the third column is the prospective pvalue and the forth column is the restrospective pvalue

rsmmat_est	<i>GLMM NULL model estimation</i>
------------	-----------------------------------

Description

This function estimate the parameters and residuals for the NULL model in SMMAT/RSMMAT test

Usage

```
rsmmat_est(y.long, time, y.cov, timecov = TRUE)
```

Arguments

y.long	Long-formatted phenotype vector
time	Time covarites matched with phenotype vector
y.cov	Covariate matrix denoting the covariate variables measured at each time
timecov	Logical variable, indicating whether the time fixed effect is estimated

Value

This function returns a list object with model parameters and residuals of the NULL GLMM model

Index

lsrat.SSD.All, [2](#)
lsrat_est, [2](#)
lsrat_simu.bi, [3](#)
lsrat_simu.conti, [3](#)
lsrat_test, [4](#)

rsmmat_est, [5](#)