Package ‘MiST’

December 13, 2018

Type  Package
Title  Set-Based Mixed Effects Score Tests for Genetic Associations between SNPs and Phenotypes
Version  0.1.0
Author  Yu-Ru Su, Li Hsu
Maintainer  Yu-Ru Su <piccoloegg@gmail.com>
Description  Functions to implement set-based tests for genetic effects of SNPs on either continuous or dichotomous outcomes under mixed effects models. It includes burden-type test allowing for external weights on each element in the set, kernel-based association test, and three tests, oMiST, aMiST and fMiST, based on combinations of the previous two tests. Set-level pvalues are produced and returned.
License  GPL (>= 2)
Depends  R (>= 3.3.2)
Imports  CompQuadForm (>= 1.4.3), stats
Encoding  UTF-8
LazyData  true
RoxygenNote  6.0.1
Suggests  knitr, rmarkdown, testthat
VignetteBuilder  knitr
NeedsCompilation  no

R topics documented:

MiST

Index

MiST

Implementation of set-based mixed effects score tests via MiST.

Description

Implementation of set-based mixed effects score tests via MiST.
**Usage**

MiST(data, d = 0, p, R = 1, outcome_type = "Continuous",
weight_method = "User", user_weight = rep(1, p), burden_scale = TRUE,
chisq_app = "3M", combinations_return = TRUE,
combination_preference = "All", acc = 5e-10, acc_auto = TRUE,
accurate_app_threshold = -log10(0.05), max_core = 4)

**Arguments**

- **data**
  A dataframe \([Y, X, G]\) with rows represent subjects and columns are \(Y\) (outcome), \(X\) (confounders), and \(G\) (genotypes).

- **d**
  Number of confounders. If \(d\) is 0, there is no confounder, and the input data should be arranged as \([Y, G]\).

- **p**
  Number of genetic variants in the set.

- **R**
  Number of burden scores, including both unweighted and weighted ones. It is always greater than or equal to 0. When \(R\) is assigned as 0, only variance component test is performed (more than one variants are required in this situation.)

- **outcome_type**
  Either "Continuous" for the values of a quantitative trait, or "Binary" for a dichotomous trait.

- **weight_method**
  One of "User" or "No weight". This specifies weights for calculating burden scores. "User" (default) allows for an equal weight across variants in the set or user-defined weights for each variant (e.g., functional annotations). The weights are specified by the input option "user_weight". "No weight" sets all weights as 0 and tests only the variance component.

- **user_weight**
  A vector or a matrix specifying an equal weight across variants or user-defined weights for calculating weighted burden score. This option only works when weight_method is set as "User". If weight_method is set as "User" and no user_weight is specified, the weight is a vector of 1.

- **burden_scale**
  A logical value (default: TRUE) specifying if standardization on burden components need to be carried out when there are more than 1 burden components.

- **chisq_app**
  Either "3M" (default) or "4M" for the moment matching (Liu’s) method in quantile approximation in optimal linear combination method. "3M" matches the 3rd moment and "4M" matches the 4th moment of the target and approximate distributions.

- **combinations_return**
  A logical value (default: TRUE) indicating if the combination methods, including the optimal linear combination, data-adaptive weighted combination and Fisher’s combination method, are implemented. By assigning FALSE, only the burden and variance component tests are conducted.

- **combination_preference**
  Either of "All" (default), or a vector containing "OptMin", "AdptWt", or "Fisher" to specify the combination method(s) to be implemented.

- **acc**
  A numerical value indicating the precision of Davies method for p-value calculation. Default is 5e-10.

- **acc_auto**
  A logical value (default: TRUE) indicating if data adaptive precision is used in optimal linear combination. We recommend to set this as TRUE for computational efficiency.
accurate_app_threshold
A numerical value specifying the threshold to determine when Liu’s and Davies
methods are used in quantile approximation in the optimal linear combination
method. Default is -log10(0.05).

max_core
An integer specifying the maximum number of cores can be recruited in parallel
package. Default is 4 cores.

Value
A list containing pvalue, stat, pvalue.ind.f, stat.ind.f, rho, and data.info. pvalue shows
p-values from the (joint) burden tests, kernel-based association test, oMiST, aMiST, and fMiST. stat
shows the corresponding statistics from the aforementioned tests. pvalue.ind.f and stat.ind.f
show p-values and test statistics of individual burden components if multiple E variables are con-
sidered jointly. rho provides weights of the burden and the kernel-based association tests in oMiST
and aMiST. data.info is a vector including p and R, the numbers of G and B terms, respectively.
Index

MiST, 1