Package ‘MiSTi’

Type Package
Title Set-Based Mixed Effects Score Tests for GxE Interactions
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 GxE interactions 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 two tests, fMiSTi and aMiSTi, based on combinations of the previous two tests. Set-level pvalues are produced and returned.
License GPL (>= 2)
Depends R (>= 3.3.2)
Imports penalized (>= 0.9.50), Matrix (>= 1.2.10), psych (>= 1.7.5), 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:

MiSTi

Index

MiSTi

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

Description

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

MiSTi(data, outcome_type = "Continuous", p, m, d = 0, R = 1,
glm_use = FALSE, weight_method = "User", user_weight = rep(1, p),
ind_burden_return = TRUE, combinations_return = TRUE,
combination_preference = "All", mac_lower_bound = 6, chisq_app = "3M",
acc = 5e-10, acc_auto = TRUE, accurate_app_threshold = -log10(0.05),
max_core = 4)

Arguments

data A nxm dataframe [Y,E,X,G] with rows represent subjects and columns are Y
(outcome), E (environmental risk factors), X (confounders), and G (genotypes).
outcome_type Either "Continuous" for the values of a quantitative trait, or "Binary" for a di-
chotomous trait.
p Number of genetic variants in the set.
m Number of environmental risk factors.
d Number of confounders. If d is 0, there is no confounder, and the input data
should be arranged as [Y E G].
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 com-
ponent test is performed (more than one GxE terms are required in this situation.)
glm_use A logical value indicating if generalized linear regression is implemented to fit
the null models. A ridge regression is implemented if glm_use is assigned as
FALSE (default).
weight_method One of "User", "No weight", "ScrnStat1", and "ScrnStat2". 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 variants (e.g., functional an-
nnotations). The weights are specified by the input option "user_weight". "No
weight" sets all weights as 0 and tests only the variance component. "ScrnStat1"
and "ScrnStat2" use screening test statistics as weights, where "ScrnStat1" use
the maximum of G-E correlation and marginal association screening test statistics
(Jiao et al. 2015) and "ScrnStat2" uses the sum of squares of marginal
association and correlation screening test statistics(Gauderman et al. 2013).
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.
ind_burden_return A logical value (default: TRUE) indicating if testing results on individual inter-
action terms between each E and burden score are returned. This option only works when
there are more than one ExB terms (i.e. multiple E's or multiple
weights for each variant).
combinations_return A logical value (default: TRUE) indicating if the combination methods, includ-
ing the optimal linear combination, data-adpative 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.
mac_lower_bound
The lower bound of minor allele count (MAC) for individual SNPs to be included in the analysis. Default is 6. SNPs with MAC \( \leq 5 \) are excluded from the analysis. A warning message is given when any SNPs are removed from the set.

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.

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 \(-\log_{10}(0.05)\).

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, oMiSTi, aMiSTi, and fMiSTi. 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 considered jointly. rho provides weights of the burden and the kernel-based association tests in oMiSTi and aMiSTi. data.info is a vector including m, p, R, pm, and Rm, the numbers of E, G, B, GxE, and BxE terms, respectively.
Index

MiSTi, 1