Package: CUMP (via r-universe)

Type Package

Title Analyze Multivariate Phenotypes by Combining Univariate Results

Version 2.0

Date 2016-08-25

Author Shuo Li `<skyli@bu.edu>`, Xuan Liu `<liuxuan@bu.edu>` and Qiong Yang `<qyang@bu.edu>`

Maintainer Shuo Li `<skyli@bu.edu>`

Description Combining Univariate Association Test Results of Multiple Phenotypes for Detecting Pleiotropy.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2016-09-15 00:27:00

Contents

CUMP-package   *Analyze Multivariate Phenotypes by Combining Univariate results*

Description

This package can be used to detect pleiotropy by combining univariate association test results of multiple phenotypes in genome-wide association studies (or studies of a large number of SNPs). In this package, we have several combination approaches including the O'Brien's method that is weighted sum of the Z or beta statistic (direction sensitive) and other methods that are weighted sum of the squared Z statistics (direction insensitive).

Details

Package:	CUMP
Type:	Package
Version:	2.0
Date:	2016-08-25
License:	GPL (>= 2)
LazyLoad:	yes

Author(s)

Shuo Li <skyli@bu.edu>, Xuan Liu <liuxuan@bu.edu> and Qiong Yang <qyang@bu.edu>

Maintainer: Shuo Li <skyli@bu.edu>

References

CUMP: an R package for analyzing multivariate phenotypes in genetic association studies

---

combGWAS	*Combining Univariate Association Test Results of Multiple Phenotypes for Detecting Pleiotropy*

Description

combGWAS() can be used to detect pleiotropy by combining univariate association test results of multiple phenotypes in genome-wide association studies (or studies of a large number of SNPs). In this function, we have several combination approaches including the O'Brien's method that is weighted sum of the Z or beta statistic (direction sensitive) and other methods that are weighted sum of the squared Z statistics (direction insensitive).

Usage

```
combGWAS(project = "mv", traitlist, traitfile, comb_method = c("z"),
betasign = rep(1, length(traitlist)), snpid, beta = NULL, SE = NULL,
Z = NULL, coded_all, AF_coded_all, n_total = NULL, pvalue = NULL,
Z_sample_weighted = FALSE)
```

Arguments

project	a character string for project name, for labeling output file.
traitlist	a vector of character strings of the phenotype names for naming the output file.

traitfile	a vector of character strings containing the univariate association results file names corresponding to the order in `traitlist` for reading in for analysis. Each univariate result file should contain the header corresponding to following fields: `snpid`, `beta`, `SE`, `coded_all`, `AF_coded_all`, `n_total`, `pvalue`. These fields can be labeled in any name in the header. You are asked to give the names of these fields as arguments for this function.
comb_method	a vector of character strings indicating the method to be used in combing univariate association results file. It can be any subset of c("z", "beta", "chisq" and "sumsq"). The details of the combination methods are given in Details.
betasign	a numeric vector for changing the signs (1 or -1) of the univariate beta (or Z) statistics. It should be of the same length and correspond to the order in `traitfile` (or `traitlist`).
snpid	the name of the genetic marker in the header of input association results files.
beta	the name of the beta estimate (if have) in the header of input association results files.
SE	the name of the standard error of the beta estimate (if have) in the header of input association results files.
Z	the name of the Z statistic (if have) in the header of input association results files.
coded_all	the name of coded allele in the header of input association results files.
AF_coded_all	the name of the allele frequency of the coded allele in the header of input association results file.
n_total	the sample size with phenotype and genotype for the genetic marker in the header of input association results file.
pvalue	the name of the p-value of the beta estimate (if have) in the header of input association results file.
Z_sample_weighted	
	a logical value. True if the results of Z method are combined by sample size weighted. False if the results are combined equally weighted.

Details

The orders of traits in `traitlist` and `traitfile` should be the same.

Currently, 4 combination methods ("z", "beta", "chisq" and "sumsq") can be implemented by the package. The default is to implement equally weighted "z" method only, but you can ask for the 4 methods simultaneously.

`betasign` should be a vector of 1 or -1 with the number of traits as the length. 1 means the beta of corresponding trait remains and -1 means the beta sign will be reversed. It only affects "z" and "beta" methods.

`snpid`, `coded_all` and `AF_coded_all` must be assigned explicitly and the corresponding columns must appear in the input datasets.

At least one of `beta`(`SE`) and `Z` should be assigned. In particular, if "beta" method is implemented,

beta and SE must be assigned.

n_total and/or pvalue can be missing in the input datasets. In particular, if "z" method is set to be sample size weighted, n_total must be assigned.

Value

No value is returned. Instead, results are written to outfile (named as "project_traits_method.csv") in the current working directory. In the outfile, there are some new variables (listed below) created by the package along with the existing variables in the original datasets.

zi	Z statistic for the ith phenotype in traitlist. They will appear in "z", "chisq" and "sumsq" methods.
pi	p-value for the ith phenotype in traitlist.
beta	combined statistic of "beta" methods.
SE	standard error of the combined statistic of "beta" methods.
Z.comb	Z statistic (Z.comb=beta/SE) of "beta" and "z" methods.
betai	beta statistic for the ith phenotype in traitlist. They will appear in output for "beta" method.
chisq.comb	combined and test statistic of "chisq" and "sumsq" methods.
pval	p-value of the combined statistic.
meanN	the mean sample size with phenotype and genotype for the genetic marker. N/A if n_total is not specified.
minN	the minimum sample size with phenotype and genotype for the genetic marker. N/A if n_total is not specified.
maxN	the maximum sample size with phenotype and genotype for the genetic marker. N/A if n_total is not specified.
remark1	The sign of beta will be flipped if coded alleles different between two datasets.
remark2	If the minimum eigen value of the covariance matrix is less than 0.01, we consider it as nearly singular and the analyses will stop.
remark3	The alleles are supposed called on positive strand. If not, the user should convert the coded allele to that on a positive strand in the result file.

Author(s)

Shuo Li <skyli@bu.edu>, Xuan Liu <liuxuan@bu.edu> and Qiong Yang <qyang@bu.edu>

References

CUMP: an R package for analyzing multivariate phenotypes in genetic association studies

Examples

```
##The following are two fake examples. Do NOT run.
##Please refer to example.pdf for details.
##no change of beta signs before combining
##combGWAS(project="mv",traitlist=c("phen1","phne2"),
## traitfile=c("Phen1GWAS.csv", "Phen2GWAS.csv"), comb_method=c("z","chisq"),
## betasign=c(1,1), snpid="SNPID", beta="beta", SE="SE",
## coded_all="coded_all"", AF_coded_all=" AF_coded_all ", pvalue="pval")

##change of  beta signs before combining: the beta sign for the 2nd phenotype reversed
##combGWAS(project="mv",traitlist=c("phen1","phne2"),
## traitfile=c("Phen1GWAS.csv", "Phen2GWAS.csv"), comb_method=c("z","chisq"),
## betasign=c(1,-1), snpid="SNPID", beta="beta", SE="SE",
## coded_all="coded_all ", AF_coded_all=" AF_coded_all ", pvalue="pval")
```

Index