Package 'rvsel'

March 19, 2014

Type Package

Title Disese/trait-related rare variant selection procedure.

Version 0.1

Date 2014-03-16

Author Hokeun Sun

Maintainer Hokeun Sun <hsun@pusan.ac.kr>

Depends Matrix

Description The most outome-related rare variants are selected within a gene or a genetic region. The selection procedure is based on the power set of the subset of the rare variants.

License GPL-2

R topics documented:

`rvsel` *Disese/trait-related rare variant selection procedure*

Description

The most outome-related rare variants are selected within a gene or a genetic region. The selection procedure is based on the power set of the subset of the rare variants.

Usage

```
rvsel(x,y,cx=NULL,family=c("gaussian","binomial"),weight=NULL,
    adapted=TRUE,ad.alpha=0.1)
```

Arguments

x	The number of genetic mutations with n samples and p varinats, where $x = 0, 1,$ or 2. It should be a n x p matrix and $p > 1$.
y	A phenotype outcome coded as 1 for cases and 0 for controls if the phenotype is case-control binary data. Otherwise, it is regarded as a quantitative outcome.
cx	Covariates such as gender and age. It should be a n x m matrix, where m is the number of covariates.
family	A type of phenotype data. "binomial" is for a case-control binary outcome and "gaussian" for a quantitative outcome. Default is "gaussian".
weight	The user defined weights for p variants. It should be the p-dimensional vector. Default is 1.
adapted	Pre-screening of potential protective variants via a marginal assoication test. If a variant is protective, the coding of x for the variant is flipped. Default is TRUE.
ad.alpha	The significant level of the margianl association test to detect potential protective variants. Default is 0.1.

Details

The weighted linear combinations of all the subset of the power set of the rare variants are generated and compared with each other to find the most outcome-related feature. When covariate effects on a phenotpye outcome exist, the residuals of linear regression with the covairates are considered to compute phenotypic association.

Adapted power set procedure first proceeds a marginal association test to detect potential protective variants where each variant is tested one at a time. If the regression coefficients from the marginal association tests are negative and the p-value of the regression coefficients is less than ad.alpha, then the variants are considered potentially protective.

Value

selection	Indicator of selected variants along with the weights of individual variants. 1 for selected variants and 0 for unselected variants.
score	Sample correlation between the weighted linear combination of the selected variants and a phenotypic outcome (residual).
jerr	The error flag

Author(s)

Hokeun Sun <hsun@pusan.ac.kr>

References

H. Sun and S. Wang (2014) *A Power Set Based Statistical Selection Procedure to Locate Susceptible Rare Variants Associated with Complex Traits with Sequencing Data*, manuscript

Examples

```
n<-2000
p<-10
MAF<-runif(p,0.001,0.01)
geno.prob<-rbind((1-MAF)^2,2*(1-MAF)*MAF,MAF^2)
```

```
x<-apply(geno.prob,2,function(x) sample(0:2,n,prob=x,replace=TRUE))
cx<-cbind(rnorm(n),sample(0:1,n,replace=TRUE))
beta<-c(rep(1,4),rep(0,6))
y<-cx%*%c(0.5,0.5)+x%*%beta+rnorm(n)
g<-rvsel(x,y,cx=cx,family="gaussian")
```

Index