Package 'IPCWK'

October 12, 2022

Title Kendall's Tau Partial Corr. for Survival Trait and Biomarkers

Version 1.0

Description We propose the inverse probability-of-censoring weighted (IPCW) Kendall's tau to measure the association of the survival trait with biomarkers and Kendall's partial correlation to reflect the relationship of the survival trait with interaction variable conditional on main effects, as described in Wang and Chen (2020) <doi:10.1093/bioinformatics/btaa017>.

Author Jie-Huei Wang, and Yi-Hau Chen

Maintainer Jie-Huei Wang <jhwang@mail.fcu.edu.tw>

Date 2020-03-26

Depends R (>= 3.4.3), MASS, survival, stats, utils

License GPL (>= 2)

Encoding UTF-8

LazyData true

Repository CRAN

RoxygenNote 6.1.1

NeedsCompilation no

Date/Publication 2020-03-26 16:20:08 UTC

R topics documented:

1

IPCWK   *Kendall's Tau Partial Corr. for Survival Trait and Biomarkers*

Description

The function proposes the inverse probability-of-censoring weighted (IPCW) Kendall's tau to measure the association of the survival trait with biomarkers and Kendall's partial correlation to reflect the relationship of the survival trait with interaction variable conditional on main effects, as described in Wang and Chen (2020) <doi:10.1093/bioinformatics/btaa017>.

Usage

```
IPCWK(data, standarize)
```

Arguments

data
> The $N \times P$ matrix of data. There are N individuals in matrix, with one individual in each row. The P columns orderly included the observable times which are time-to-event or censoring times and without ties at the event times, the status is a binary variable with 1 indicating the event has occured and 0 indicating (right) censoring, and the $(P - 2)$ main predictors. Note that the missing values of predictors in the data are not allowed in this version.

standarize
> Setting to "FALSE", point out the original gene espressions profiles are not standarized. We are going to standarize the gene features automatically by IPCWK function. Setting to "TRUE", point out the original gene espressions profiles have been standarized. We maintain the original gene features matrix to do the following analysis.

Value

Returns a list with components

CORR
> The $3 \times K$ proposed correlation matrix, where K is the dimension of all biomarkers include main and interaction effects with one method in each row. The first row means the IPCW Kendall's tau correlation; the second row means the original Kendall's tau correlation without considering censoring scheme and the final row means the proposed Kendall's partial correlation. Note that the column names of "CORR" matrix is gene index, in which "Ga" means the ath main gene and "Ga&Gb" means the ath main gene interact with the bth main gene.

References

Wang JH, and Chen YH* (2020) Interaction Screening by Kendall's Partial Correlation for Ultrahigh-dimensional Data with Survival Trait. published in Bioinformatics <doi:10.1093/bioinformatics/btaa017>.

Examples

```
set.seed(123)
library(MASS)
library(survival)

numbeta=50
N=100
beta0=matrix(0, numbeta, 1)
sigma1=0.5^abs(outer(1:numbeta, 1:numbeta, "-"))
W1=mvrnorm(N, beta0, sigma1, tol=1e-8, empirical=FALSE)
Z1=W1^2-1

### produce quadratic and two-way interaction effects ###
tempZZ1=matrix(Z1^2, N, numbeta)
tempZZ2=model.matrix(~(.)^2 - . -1, data=data.frame(Z1))
AZ=cbind(Z1, cbind(tempZZ1, tempZZ2))

### identify true predictors: G1, G10, G1&G1, G10&G10, G1&G10, G10&G20 ###
a=3
beta=matrix(0, dim(AZ)[2], 1)
beta[1,]=-0.8*a
beta[10,]=a
beta[51,]=1.2*a
beta[60,]=a
beta[109,]=-1.2*a
beta[515,]=a

### simulate survival time follows linear transformation model ###
C=matrix(runif(N,0,1), N, 1)
ST=X=S=matrix(0, N, 1)
temp=rexp(N)
ST=as.numeric(0.5*log(2*temp*exp(-AZ%*%beta)+1))
X=pmin(ST, C)
S=(ST==X)
survdata=cbind(X, S, Z1)

### perform IPCWK function ###
test=IPCWK(data=survdata, standarize="FALSE")
true=which(beta!=0)
sum(order(-abs(test$CORR[1,]))[1:20] %in% true) ### IPCW-tau
sum(order(-abs(test$CORR[2,]))[1:20] %in% true) ### Kendall's tau
sum(order(-abs(test$CORR[3,]))[1:20] %in% true) ### PC-IPCW-tau
```

Index