Package ‘ctgt’

June 7, 2021

Type Package
Title Closed Testing with Globaltest for Pathway Analysis
Version 2.0
Date 2021-06-07
Author Ningning Xu
Maintainer Ningning Xu <n.xu@lumc.nl>
Description A shortcut procedure is proposed to implement closed testing for large-scale multiple testings, especially with the global test. This shortcut is asymptotically equivalent to closed testing and post hoc. Users could detect any possible sets of features or pathways with family-wise error rate controlled. The global test is powerful to detect associations between a group of features and an outcome of interest.
License GPL (>= 2)
Imports Rcpp (>= 1.0.3)
LinkingTo Rcpp, BH
SystemRequirements C++11
NeedsCompilation yes
Repository CRAN
Date/Publication 2021-06-07 15:10:02 UTC

R topics documented:

tgt-package ................................................................. 2
l-alpha quantile ............................................................ 2
closed testing with globaltest ........................................... 3
internalfunctions .......................................................... 4
majorizing vector ......................................................... 6
non-standardized globaltest .............................................. 6
p-value ..................................................................... 8
true discoveries .............................................................. 8

Index 10
Description

A shortcut procedure for closed testing with the global test is presented.

Details

See examples in `actgt` function.

Author(s)

Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References


1-alpha quantile

The 1-alpha quantile of globaltest

Description

Robbins and Pitman Algorithm to calculate the critical value given eigenvalue vector and alpha level.

Usage

```r
criticalvalue(lam, alpha = 0.05)
```

Arguments

- `lam` The numeric vector with eigenvalues as elements.
- `alpha` The type I error rate allowed. The default is 0.05.

Value

Returns a real number.
closed testing with globaltest

Author(s)
Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References

closed testing with globaltest

Approximated Closed Testing with Global Test

Description
To detect the significance of the set of features after correcting for multiple global tests, with family-wise error rate controlled.

Usage
actgt (y, X, xs, hyps, maxit = 0, alpha = 0.05)

Arguments
- y: The response vector (numeric vector).
- X: The full design matrix, whose columns are named by the covariates.
- xs: The name vector of all covariates (character vector).
- hyps: The name vector of the covariates in the pathway of interest (character vector).
- maxit: An optional integer to denote the maximal iterations for branch and bound method. The default value 0 means the single-step shortcut without branch and bound method. Note that larger value is more time-consuming.
- alpha: The type I error rate allowed. The default is 0.05.

Value
Returns a list of rejection indicator and the number of iterations.

Author(s)
Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References
Examples

# Generate the design matrix and response vector for logistic regression models
n = 100
m = 5
X = matrix(data = 0, nrow = n, ncol = m, byrow = TRUE)
for (i in 1:n)
    set.seed(1234+i)
    X[i,] = as.vector(arima.sim(model = list(order = c(1, 0, 0), ar = 0.2), n = m))

y = rbinom(n, 1, 0.6)
X[which(y==1),1:3] = X[which(y==1),1:3] + 0.8

xs = paste("x", seq(1, m, 1), sep="")
colnames(X) = xs

hyps = xs[1]

# The single-step ctgt procedure
actgt(y = y, X = X, xs = xs, hyps = hyps, maxit = 0, alpha = 0.05)
# Result Iterations
# "unsure" "0"

# The iterative ctgt procedure with more iterations
actgt(y = y, X = X, xs = xs, hyps = hyps, maxit = 0, alpha = 0.05)
# Result Iterations
# "reject" "2"
# which means that x1 is rejected by closed testing within two more iterations of the shortcut

# For a group of feature sets
mysets = list(xs[1:5], xs[c(1, 4)], xs[c(1, 4, 5)])
sapply(mysets, function(i) actgt(y = y, X = X, xs = xs, hyps = i, maxit = 0, alpha = 0.05))
# Result "reject" "unsure" "reject"
# Iterations "0" "0" "0"

mysets = list(xs[1:5], xs[c(1, 4)], xs[c(1, 4, 5)])
sapply(mysets, function(i) actgt(y = y, X = X, xs = xs, hyps = i, maxit = 0, alpha = 0.05))
# Result "reject" "reject" "reject"
# Iterations "0" "2" "0"

internalfunctions

'Internal Functions (ctgt)'

Description

Internal functions of ctgt.
Usage

```r
## iterative shortcut with branch and bound
actgt_it(y, Tmatrix, Cmatrix, fxs, sxs, Tf, Lamf, Cf, Ts, Lams, Cs, count=1, maxIt=1, a = 0.05)

## to check whether tmin is above cmax
tacmax(tmins, levels, tw, cf, lf, ls, alp)

## to check whether tmin is above ctrue
tactrue(tmins, hyxs, cfull, Wmatrix, alp)
```

Arguments

- `y` The response vector (numeric vector).
- `Tmatrix` The matrix used to calculate the test statistics.
- `Cmatrix` The matrix used to calculate the critical values.
- `fxs` The name vector of upper model (character vector).
- `sxs` The name vector of lower model (character vector).
- `Tf, Lamf, Cf` Test statistic, eigenvalues, and critical value of fxs.
- `Ts, Lams, Cs` Test statistic, eigenvalues, and critical value of sxs.
- `count` count the branches, default is 1.
- `maxIt` maximal number of branches chosen by user, default is 1.
- `a, alp` alpha level.
- `tmins` Minimum test statistics.
- `levels` levels
- `tw, cf, lf, ls` sorted weights, critical values and level for fxs and sxs.
- `hyxs` The name vector of the covariates of interest (character vector).
- `cfull` critical value of full model.
- `Wmatrix` matrix to calculate majorizing vector.

Author(s)

Ningning Xu

Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>
The majorizing vector

Description
To get the majorizing vector at a specific level, given the upper bound and lower bound.

Usage
getL (ub, lb, level)

Arguments
- ub: upper bound.
- lb: lower bound.
- level: level of interest.

Value
Returns a numeric vector with the same length as ub and lb.

Author(s)
Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References

'Non-standardized globaltest'

Description
This is the second version of the globaltest, the non-standardized globaltest

Usage
## a powerful variant of globaltest
gt2 (y, X, hyps, alpha = 0.05)
non-standardized globaltest

Arguments

y The response vector (numeric vector).
X The full design matrix, whose columns are named by the covariates.
hyps The name vector of the covariates in the pathway of interest (character vector).
alpha The type I error rate allowed. The default is 0.05.

Value

Returns the p-value, the observed and expected test statistics and the number of covariates.

Author(s)

Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References


Examples

#Generate the design matrix and response vector for logistic regression models
n = 100
m = 5
X = matrix(data = 0, nrow = n, ncol = m, byrow = TRUE)
for (i in 1:n){
  set.seed(1234+i)
  X[i,] = as.vector(arima.sim(model = list(order = c(1, 0, 0), ar = 0.2), n = m))
}
y = rbinom(n, 1, 0.6)
X[which(y==1),1:3] = X[which(y==1),1:3] + 0.8

xs = paste("x",seq(1,m,1),sep="")
colnames(X) = xs

hyps=xs[1]

#The raw p-values of globaltest
gt2(y = y, X = X, hyps = hyps, alpha = 0.05)
#p-value Statistic Expected    #Cov
#7.64e-03  2.30e+02  1.24e+02    1.00e+00
### Robbins and Pitman Algorithm

Robbins and Pitman Algorithm to calculate the p-value given the observed value and the eigenvalue vector.

#### Usage

\[
pv(x, \text{lam})
\]

#### Arguments

- **x**: The observed value that is used to calculate the corresponding p-value.
- **lam**: The numeric vector with eigenvalues as elements.

#### Value

Returns a value between 0 and 1.

#### Author(s)

Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

#### References


---

### True Discoveries

To count the number of true discoveries within a given pathway or feature set of interest.

#### Usage

\[
discoveries(y, X, xs, hyps, maxit = 0, alpha = 0.05)
\]
true discoveries

Arguments

- **y**: The response vector (numeric vector).
- **X**: The full design matrix, whose columns are named by the covariates.
- **xs**: The name vector of all covariates (character vector).
- **hyps**: The name vector of the covariates in the pathway of interest (character vector).
- **maxit**: An optional integer to denote the maximal iterations for branch and bound method. The default value 0 means the single-step shortcut without branch and bound method. Note that larger value is more time-consuming.
- **alpha**: The type I error rate allowed. The default is 0.05.

Value

Returns a non-negative integer.

Author(s)

Ningning Xu
Maintainer: Ningning Xu <n.xu@lumc.nl; xu15263142750@gmail.com>

References


Examples

```r
#Generate the design matrix and response vector for logistic regression models
n = 100
m = 5
X = matrix(data = 0, nrow = n, ncol = m, byrow = TRUE)
for (i in 1:n)
  set.seed(1234+i)
  X[i,] = as.vector(arima.sim(model = list(order = c(1, 0, 0), ar = 0.2), n = m))

y = rbinom(n, 1, 0.6)
X[which(y==1),1:3] = X[which(y==1),1:3] + 0.8
xs = paste("x", seq(1, m, 1), sep = "")
colnames(X) = xs

# For standarized data
X = scale(x = X, center = FALSE, scale = TRUE)/sqrt(n-1)
interest = xs
discoveries(y=y, X = X, xs = xs, hyps = interest)
#2
discoveries(y=y, X = X, xs = xs, hyps = interest, maxit=10)
#2
```
Index

* package
  ctgt-package, 2
  1-alpha quantile, 2

actgt, 2
actgt (closed testing with globaltest), 3
actgt_it (internalfunctions), 4

closed testing with globaltest, 3
criticalvalue (1-alpha quantile), 2
ctgt (ctgt-package), 2
ctgt-package, 2

discoveries (true discoveries), 8

ggetL (majorizing vector), 6
gt2 (non-standardized globaltest), 6

internalfunctions, 4

majorizing vector, 6

non-standardized globaltest, 6

p-value, 8
pv (p-value), 8

tacmax (internalfunctions), 4
tactrue (internalfunctions), 4
true discoveries, 8