Package ‘gtcorr’

February 20, 2015

Type Package

Title Calculate efficiencies of group testing algorithms with correlated responses

Version 0.2-1

Date 2011-05-24

Author Sam Lendle <`lendle@stat.berkeley.edu`>

Maintainer Sam Lendle <`lendle@stat.berkeley.edu`>

Description This package provides functions to calculate the efficiencies (expected tests per unit) of hierarchical and matrix group testing procedures. Efficiencies can be calculated in the presense of correlated responses under multiple arrangements of clustesrs. Efficiencies can also be evaluated in the presense of test error.

License GPL-3

LazyLoad yes

Repository CRAN

Date/Publication 2011-05-26 20:48:48

NeedsCompilation no

R topics documented:

`gtcorr-package` *Calculate efficiencies of group testing algorithms with correlated responses*

Description

This package provides functions to calculate the efficiencies (expected tests per unit) of hierarchical and matrix group testing procedures. Efficiencies can be calculated in the presence of correlated responses under multiple arrangements of clusters. Efficiencies can also be evaluated in the presence of test error.

Author(s)

Sam Lendle <lendle@stat.berkeley.edu>

References

Samuel D. Lendle, Michael Hudgens, and Bahjat F. Qaqish, "Group Testing for Case Identification with Correlated Responses" Submitted 2011. Biometrics.

`gtcorr.hierarchical` *Calculate the efficiency of hierarchical group testing procedures for nested and random arrangements*

Description

'gtcorr.hierarchical' calculates the efficiencies of hierarchical group testing procedures for nested and random arrangements, allowing for correlation between units and test error.

Usage

```
gtcorr.hierarchical(n, m = 1, p, sigma = 0, se = 1, sp = 1, arrangement
= c("nested", "random"), model = c("beta-binomial", "Madsen",
"Morel-Neerchal"), ...)
```

Arguments

`n`	a numeric vector of pool sizes where `n[s]` is the size of a pool in stage `s`. The size of a pool in the last stage is 1, which can be omitted.
`m`	cluster size.
`p`	probability of a unit testing positive.
`sigma`	pairwise correlation of two units in a cluster.
`se`	sensitivity. The probability that a pool of units tests positive given than at least one unit in that pool is positive

`sp`	specificity. The probability that a pool of units tests negative given that at least one unit in that pool is negative
`arrangement`	how clusters are arranged. Should be ‘nested’ or ‘random’.
`model`	probability model for clusters. Should be ‘beta-binomial’, ‘Madsen’, or ‘Morel-Neerchal’.
`...`	`runs` for a random arrangement, number of Monte Carlo simulations to perform to calculate the probability of a pool having no positive units. Default is 1000.

Details

One of `m`, `p`, `sigma`, `se`, or `sp` can have more than one value. `m` should not be greater than `n[1]`. For a ‘nested’ arrangement, `m` should be divisible by `n[s]` or `n[s]` should be divisible by `m` for all `s`. See Lendle et. al. 2011 for more information.

Value

`n`	number of units per pool at each stage.
`param.grid`	a data frame containing the values of `p`, `sigma`, `se`, `sp`, and `m` for each value of `efficiency`.
`arrangement`	arrangement.
`model`	model.
`efficiency`	a vector of efficiencies, one for each row of `param.grid`.

References

Samuel D. Lendle, Michael Hudgens, and Bahjat F. Qaqish, "Group Testing for Case Identification with Correlated Responses" Submitted 2011. Biometrics.

See Also

`gtcorr.hierarchical.user`

Examples

```
##Plot efficiencies of a Dorfman (2 stage hierarchical) algorithm
##by cluster size and sigma
m <- 2^(0:8)
sig.0 <- gtcorr.hierarchical(n=256, p=.001, m=m, sigma=0)$efficiency
sig.05 <- gtcorr.hierarchical(n=256, p=.001, m=m, sigma=0.05)$efficiency
sig.5 <- gtcorr.hierarchical(n=256, p=.001, m=m, sigma=0.5)$efficiency
sig.99 <- gtcorr.hierarchical(n=256, p=.001, m=m, sigma=.99)$efficiency

plot(m, sig.99, log=”x”, type='b', ylab=”Efficiency”, axes=FALSE)
box()
axis(1, at=m)
axis(2)
lines(m, sig.5, type='b', pch=22)
lines(m, sig.05, type='b', pch=23)
```

```
lines(m, sig.0, type='b', pch=24)
legend('bottomleft', c("sigma=0", "sigma=0.05", "sigma=0.5",
"sigma=0.99"), pch=21:24)
```

`gtcorr.hierarchical.user`

Calculate the efficiency of hierarchical group testing procedures for user specified arrangements

---

Description

'gtcorr.hierarchical.user' calculates the efficiencies of hierarchical group testing procedures, allowing for correlation between units and test error. Cluster arrangements can are specified by the user and probability of testing postive and correlation can vary by cluster.

Usage

```
gtcorr.hierarchical.user(n, clusters, p, sigma=0, se=1, sp=1, model =
c("beta-binomial", "Morel-Neerchal", "Madsen"))
```

Arguments

`n`	a numeric vector of pool sizes where `n[s]` is the size of a pool in stage `s`. The size of a pool in the last stage is 1, which can be omitted.
`clusters`	a vector of length `n[1]` of integers from 1 up to the total number of clusters. `cluster[i]` is the ith unit's cluster.
`p`	probability of a unit testing positive. If the length is one, then all clusters have the same probability of testing positive. If the length is the total number of clusters, then `p[k]` is the probability that a unit in the kth cluster tests positive.
`sigma`	pairwise correlation of two units in a cluster. If the length is one, then all clusters have the same pairwise correlation. If the length is the total number of clusters, then `sigma[k]` is the pairwise correlation for the kth cluster.
`se`	sensitivity. The probability that a pool of units tests positive given than at least one unit in that pool is positive
`sp`	specificity. The probability that a pool of units tests negative given that at least one unit in that pool is negative
`model`	probability model for clusters. Should be 'beta-binomial', 'Madsen', or 'Morel-Neerchal'.

Details

Units are ordered such that the first pool in stage s contains units `1, ..., n[s]`, and the second pool in stage s has units `(n[s]+1),...,(2*n[s])` and so forth.

Value

A length one vector containing the efficiency is returned

References

Samuel D. Lendle, Michael Hudgens, and Bahjat F. Qaqish, "Group Testing for Case Identification with Correlated Responses" Submitted 2011. Biometrics.

See Also

`gtcorr.hierarchical`

Examples

```
##Calculate the efficiency of a two stage hierarchical procedure with
##n[1]=20, where the first cluster has 10 units and the second and third
##clusters have 5 units each.
n <- 20
clusters <- c(rep(1,10), rep(2, 5), rep(3, 5))
p <- .1
sigma <- .3
gtcorr.hierarchical.user(n, clusters, p, sigma)
```

`gtcorr.matrix` *Calculate the efficiency of matrix group testing procedures for rectangular, diagonal, and random arrangements*

Description

'gtcorr.matrix' calculates the efficiencies of matrix group testing procedures for rectangular, diagonal, and random arrangements, allowing for correlation between units and test error.

Usage

```
gtcorr.matrix(r, c, m = 1, p, sigma = 0, se = 1, sp = 1, r.prime, c.prime, arrangement = c("rectangular
```

Arguments

`r`	number of rows in the pooling matrix.
`c`	number of columns in the pooling matrix.
`m`	cluster size.
`p`	probability of a unit testing positive.
`sigma`	pairwise correlation of two units in a cluster.
`se`	sensitivity. The probability that a pool of units tests positive given than at least one unit in that pool is positive

`sp`	specificity. The probability that a pool of units tests negative given that at least one unit in that pool is negative
`r.prime`	for a 'rectangular' arrangement, the number of rows in a rectangular cluster.
`c.prime`	for a 'rectangular' arrangement, the number of columns in a rectangular cluster.
`arrangement`	how clusters are arranged. Should be 'rectangular', 'diagonal' or 'random'.
`model`	probability model for clusters. Should be 'beta-binomial', 'Madsen', or 'Morel-Neerchal'.
`...`	`runs` for a random arrangement, number of Monte Carlo simulations to perform to calculate the probability of a pool having no positive units. Default is 1000.

Details

One of `m`, `p`, `sigma`, `se`, or `sp` can have more than one value. For a diagonal arrangement, `r`, `c`, and `m` should be equal. For a rectangular arrangement, `m` should be `r.prime*c.prime`. See Lendle et. al. 2011 for more information.

Value

`r`	
`c`	
`m`	cluster size.
`r.prime`	number of rows in the pooling matrix.
`c.prime`	number of columns in the pooling matrix.
`param.grid`	a data frame containing the values of `p`, `sigma`, `se`, and `sp` for each value of `efficiency`.
`arrangement`	arrangement.
`model`	model.
`efficiency`	a vector of efficiencies, one for each row of `param.grid`.

References

Samuel D. Lendle, Michael Hudgens, and Bahjat F. Qaqish, "Group Testing for Case Identification with Correlated Responses" Submitted 2011. Biometrics.

See Also

`gtcorr.matrix.user`

Examples

```
##Plot efficiencies of a 16 by 16 matrix procedure by arrangement
sigma <- seq(0,.99, length.out=100)
sig2 <- seq(0, .99, length.out=10)
diag <- gtcorr.matrix(r=16, c=16, m=16, r.prime=1, c.prime=16,
                      arr='diag', p=.05, sigma=sigma)$efficiency
```

```
rand <- gtcorr.matrix(r=16, c=16, m=16, r.prime=1, c.prime=16,
                      arr='rand', p=.05, sigma=sig2)$efficiency
rect1 <- gtcorr.matrix(r=16, c=16, m=16, r.prime=1, c.prime=16, p=.05,
                       sigma=sigma)$efficiency
rect2 <- gtcorr.matrix(r=16, c=16, m=16, r.prime=2, c.prime=8, p=.05,
                       sigma=sigma)$efficiency
rect3 <- gtcorr.matrix(r=16, c=16, m=16, r.prime=4, c.prime=4, p=.05,
                       sigma=sigma)$efficiency

plot(sigma, diag, ylim=c(0, max(diag)), type='l', ylab="Efficiency", xlab="sigma")
lines(sig2, rand, col=2)
lines(sigma, rect3, col=3)
lines(sigma, rect2, col=4)
lines(sigma, rect1, col=5)
legend("bottomleft", c("Diagonal", "Random", "4x4 rect.", "2x8 rect.",
                       "1x16 rect."), lty=1, col=1:5)
```

`gtcorr.matrix.user` *Calculate the efficiency of matrix group testing procedures for user specified arrangements*

Description

'gtcorr.matrix.user' calculates the efficiencies of matrix based group testing procedures, allowing for correlation between units and test error. Cluster arrangements can are specified by the user and probability of testing postive and correlation can vary by cluster.

Usage

```
gtcorr.matrix.user(clusters, p, sigma=0, se=1, sp=1, model = c("beta-binomial", "Madsen", "Morel-Neer
```

Arguments

`clusters`	a matrix of integers from 1 up to the total number of clusters. `cluster[i,j]` is the ith, jth unit's cluster.
`p`	probability of a unit testing positive. If the length is one, then all clusters have the same probability of testing positive. If the length is the total number of clusters, then `p[k]` is the probability that a unit in the kth cluster tests positive.
`sigma`	pairwise correlation of two units in a cluster. If the length is one, then all clusters have the same pairwise correlation. If the length is the total number of clusters, then `sigma[k]` is the pairwise correlation for the kth cluster.
`se`	sensitivity. The probability that a pool of units tests positive given than at least one unit in that pool is positive
`sp`	specificity. The probability that a pool of units tests negative given that at least one unit in that pool is negative
`model`	probability model for clusters. Should be 'beta-binomial', 'Madsen', or 'Morel-Neerchal'.

Value

A length one vector containing the efficiency is returned

References

Samuel D. Lendle, Michael Hudgens, and Bahjat F. Qaqish, "Group Testing for Case Identification with Correlated Responses" Submitted 2011. Biometrics.

See Also

`gtcorr.matrix`

Examples

```
##Calculate the efficiency of a 4 by 4 matrix procedure where the first
##two rows are in the first cluster, the bottom left 2 by 2 submatrix
##makes up the second cluster, and the bottom right 2 by 2 submatrix
##makes up the third cluster.
clusters <- matrix(NA, 4, 4)
clusters[1:2, ] <- 1
clusters[3:4, 1:2] <- 2
clusters[3:4, 3:4] <- 3
p <- .01
sigma <- .4
gtcorr.matrix.user(clusters, p, sigma)
```

Index