Package ‘MHThier’

February 6, 2017

Type Package

Title Multiple Hypotheses Testing for Hierarchical Data

Version 0.1.0

Imports igraph, methods, phyloseq, ape

Suggests structSSI, multtest, FixSeqMTP

Author Yalin Zhu, Wenge Guo, Gavin Lynch

Maintainer Yalin Zhu <yalin.zhu@outlook.com>

BugReports https://github.com/allenzhuaz/MHThier/issues

Description This package performs several multiple testing procedures that consider hierarchical structure among the tested hypotheses. The generalized hierarchical procedures control the FDR under various forms of dependence structure such as positive dependence and block dependence. Our functions provide an option to choose the type of dependence.

License GPL (>= 2)

Encoding UTF-8

LazyData true

R topics documented:

`ArbitraryDeptCF` *Critical Function for Hierarchical FDR Controlling Procedure under Arbitrary Dependence.*

Description

Given some parameters of the descendant hypotheses and ancestor hypotheses, return the critical function using the generalized hierarchical FDR controlling procedure under arbitrary dependence (see Theorem 2 in Lynch et al. (2016)).

Usage

```
ArbitraryDeptCF(isTested, mi, li, l, depth, gdi, alpha, rOffset)
```

Arguments

`isTested`	logical; if `TRUE` (default), then the i-th hypotheses H_i will be tested; otherwise, H_i will not be tested.
`mi`	the cardinality of the set of descendant hypotheses of H_i.
`li`	the number of leaf hypotheses in the set of descendant hypotheses H_i.
`l`	the total number of leaf hypotheses.
`depth`	the cardinality of the set of ancestor hypotheses of H_i, referred to as the depth of H_i.
`gdi`	the cardinality of the union set of all hypotheses with all depth no deeper than the given depth.
`alpha`	the significant level used to calculate the critical values to make decisions.
`rOffset`	the offset increment for the number of rejections.

Value

A critical function of the index i and rejection number R.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

`BlockArbitraryCF` *Critical Function for Hierarchical FDR Controlling Procedure under Block Arbitrary Dependence.*

Description

Given some parameters of the descendant hypotheses and ancestor hypotheses, return the critical function using the generalized hierarchical FDR controlling procedure under block arbitrary dependence (see Theorem 4 in Lynch et al. (2016)).

Usage

```
BlockArbitraryCF(isTested, mi, li, l, depth, fdi, alpha, rOffset)
```

Arguments

`isTested`	logical; if `TRUE` (default), then the i-th hypotheses H_i will be tested; otherwise, H_i will not be tested.
`mi`	the cardinality of the set of descendant hypotheses of H_i.
`li`	the number of leaf hypotheses in the set of descendant hypotheses H_i.
`l`	the total number of leaf hypotheses.
`depth`	the cardinality of the set of ancestor hypotheses of H_i, referred to as the depth of H_i.
`fdi`	the cardinality of the set of all hypotheses with the given depth.
`alpha`	the significant level used to calculate the critical values to make decisions.
`rOffset`	the offset increment for the number of rejections.

Value

A critical function of the index i and rejection number R.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

`BlockPositiveCF` *Critical Function for Hierarchical FDR Controlling Procedure under Block Positive Dependence*

Description

Given some parameters of the descendant hypotheses, return the critical function using the generalized hierarchical FDR controlling procedure under block positive dependence (see Theorem 3 in Lynch et al. (2016)).

Usage

```
BlockPositiveCF(isTested, mi, li, l, alpha, rOffset)
```

Arguments

`isTested`	logical; if `TRUE` (default), then the i-th hypotheses H_i will be tested; otherwise, H_i will not be tested.
`mi`	the cardinality of the set of descendant hypotheses H_i.
`li`	the number of leaf hypotheses in the set of descendant hypotheses H_i.
`l`	the total number of leaf hypotheses.
`alpha`	the significant level used to calculate the critical values to make decisions.
`rOffset`	the offset increment for the number of rejections.

Value

A critical function of the index i and rejection number R.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

`hier.test` *Decision-making Function for Generalized Hierarchical Testing Procedures*

Description

Given a set of p-values, return the decisions using the generalized stepwise procedure.

Usage

```
hier.test(tree, pvals, alpha, type)
```

Arguments

`tree`	the edgelist parameterizing the hierarchical structure between hypotheses. The edges must be stored so that each edge is a row of a two column matrix, where the first column gives the parent and the second gives the child.
`pvals`	a vector of raw p-values resulting from an experiment. The names of this vector should be contained in the edgelist parameterizing the hierarchical structure between hypothesis, inputted as `tree`
`alpha`	the significant level used to calculate the critical values to make decisions.
`type`	the type of dependence structure of the hierarchically ordered hypotheses. Currently, we provide four types of dependence: `"positive"`, `"arbitrary"`, `"block positive"` and `"block arbitrary"`.

Value

logical values of each hypothesis being rejected or not, if `TRUE`, then the hypothesis is rejected; otherwise, the hypothesis is not rejected.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

Examples

```
library(igraph)
library(ape)
library(structSSI)
library(phyloseq)
data("chlamydiae")
environments <- sample_data(chlamydiae)$SampleType
abundances <- otu_table(chlamydiae)

graph.tree <- as.igraph(phy_tree(chlamydiae))
edge.tree <- get.edgelist(graph.tree)
pVal <- treePValues(edge.tree, abundances, environments)
pVal[which(is.na(pVal))] = 1; # these have all 0 abundances in every environment
decision1 <- hier.test(tree = graph.tree, pvals = pVal, alpha = 0.01, type = "positive")
decision2 <- hier.test(tree = graph.tree, pvals = pVal, alpha = 0.01, type = "arbitrary")
## show the number of rejections under different types of dependence
length(which(decision1 == TRUE)); length(which(decision2 == TRUE))
```

`hierFDR.CF`	*Critical Function for all Hierarchical FDR Controlling Procedure under Various Types of Dependence.*

Description

Given some parameters of the descendant hypotheses and ancestor hypotheses, return the critical function using the generalized hierarchical FDR controlling procedure under various type of dependence.

Usage

```
hierFDR.CF(isTested, mi, li, l, depth, fdi, gdi, alpha, rOffset, type)
```

Arguments

`isTested`	logical; if `TRUE` (default), then the i-th hypotheses H_i will be tested; otherwise, H_i will not be tested.
`mi`	the cardinality of the set of descendant hypotheses of H_i.
`li`	the number of leaf hypotheses in the set of descendant hypotheses H_i.
`l`	the total number of leaf hypotheses.
`depth`	the cardinality of the set of ancestor hypotheses of H_i, referred to as the depth of H_i.
`fdi`	the cardinality of the set of all hypotheses with the given depth.
`gdi`	the cardinality of the union set of all hypotheses with all depth no deeper than the given depth.
`alpha`	the significant level used to calculate the critical values to make decisions.
`rOffset`	the offset increment for the number of rejections.
`type`	the type of dependence structure of the hierarchically ordered hypotheses. Currently, we provide four types of dependence: `"positive"`, `"arbitrary"`, `"block positive"` and `"block arbitrary"`.

Value

A critical function of the index i and rejection number R.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

See Also

`PositiveDeptCF`, `ArbitraryDeptCF`, `BlockPositiveCF`, `BlockArbitraryCF`

`PositiveDeptCF` *Critical Function for Hierarchical FDR Controlling Procedure under Positive Dependence*

Description

Given some parameters of the descendant hypotheses, return the critical function using the generalized hierarchical FDR controlling procedure under positive dependence (see Theorem 1 in Lynch et al. (2016)).

Usage

```
PositiveDeptCF(isTested, mi, li, l, alpha, rOffset)
```

Arguments

`isTested`	logical; if `TRUE` (default), then the i-th hypotheses H_i will be tested; otherwise, H_i will not be tested.
`mi`	the cardinality of the set of descendant hypotheses H_i.
`li`	the number of leaf hypotheses in the set of descendant hypotheses H_i.
`l`	the total number of leaf hypotheses.
`alpha`	the significant level used to calculate the critical values to make decisions.
`rOffset`	the offset increment for the number of rejections.

Value

A critical function of the index i and rejection number R.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

`stepwise` *Decision-making Function for Generalized Stepwise Procedures*

Description

Given a set of p-values, return the decisions using the generalized stepwise procedure.

Usage

```
stepwise(p, CF, k)
```

Arguments

p	numeric vector of p-values (possibly with NAs). Any other R is coerced by as.numeric. Same as in p.adjust.
CF	critical functions of the corresponding multiple testing procedure
k	the order of a generalized stepwise procedure.

Value

logical values of each hypothesis being rejected or not, if TRUE, then the hypothesis is rejected; otherwise, the hypothesis is not rejected.

Author(s)

Yalin Zhu

References

Lynch, G., Guo, W. (2016). On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses. *arXiv preprint* arXiv:1612.04467.

Tamhane, A. C., Liu, W., & Dunnett, C. W. (1998). A generalized step-up-down multiple test procedure. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 353-363.

Sarkar, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Annals of statistics*, 239-257.

Index