

# Package ‘discreteMTP’

February 19, 2015

**Type** Package

**Title** Multiple testing procedures for discrete test statistics

**Version** 0.1-2

**Date** 2012-05-04

**Description** Multiple testing procedures for discrete test statistics, that use the known discrete null distribution of the p-values for simultaneous inference.

**License** GPL (>= 2)

**URL** <http://www.math.tau.ac.il/~ruheller/>

**Author** Ruth Heller [aut], Hadas Gur [aut], Shay Yaacoby [aut, cre]

**Maintainer** Shay Yaacoby <shay66@gmail.com>

**Repository** CRAN

**Date/Publication** 2012-05-05 09:25:07

**NeedsCompilation** no

## R topics documented:

discreteMTP-package . . . . .	1
amnesia . . . . .	2
p.discrete.adjust . . . . .	3

<b>Index</b>	<b>6</b>
--------------	----------

---

discreteMTP-package    *Multiple Testing Procedures For Discrete Test Statistics*

---

## Description

Multiple testing procedures for discrete test statistics, that use the known discrete null distribution of the p-values for simultaneous inference. The package contains a `p.discrete.adjust` function, similar to `p.adjust`, that provides the adjusted p-values from two step-up and two step-down multiple testing procedures that exploit the discrete null distribution of the p-values. A detailed description of these procedures can be found in Heller & Gur (2011) [link](#).

**Details**

Package: discreteMTP  
Type: Package  
Version: 0.1.1  
Date: 2012-04-28  
License: GPL (>= 2)

**Author(s)**

Ruth Heller, Hadas Gur and Shay Yaacoby.

Maintainer: Shay Yaacoby <shay66@gmail.com>.

---

amnesia

*Amnesia And Other Drug Reactions In The MHRA Pharmacovigilance  
Spontaneous Reporting System*

---

**Description**

For each of 2446 drugs in the MHRA database (column 1), the number of cases with Amnesia as an adverse event (column 2), and the number of cases with adverse event for this drug (column 3). In total, 682648 adverse drug reactions were reported, among them 2044 cases of amnesia.

**Usage**

```
data(amnesia)
```

**Format**

A data frame with 2446 rows representing drugs with the following 3 columns:

DrugName The name of the drug.

AmnesiaCases Number of the amnesia cases reported for the drug.

AllAdverseCases Number of all adverse drug reactions reported for the drug.

**Details**

The data was collected from the Drug Analysis Prints published by the Medicines and Healthcare products Regulatory Agency (MHRA).

**Source**

[Drug Analysis Prints on MHRA site.](#)

## References

Heller, R., and Gur, H. (2011). False discovery rate controlling procedures for discrete tests. *arXiv:1112.4627v1* [link](#).

---

p.discrete.adjust      *Adjusted Discrete Distributed P-values for Multiple Testing*

---

## Description

Given a set of p-values and their discrete cumulative distribution functions (CDF), returns p-values adjusted using one of several methods.

## Usage

```
p.discrete.adjust(p, pCDF, method = p.discrete.adjust.methods, cutoff = 1, n = length(p))
```

```
p.discrete.adjust.methods
## c("BH", "BL", "BHmidp", "BLmidp", "DBH", "DBL", "none")
```

## Arguments

p	numeric vector of p-values (possibly with NAs). Any other R is coerced by <a href="#">as.numeric</a> . Same as in <a href="#">p.adjust</a> .
pCDF	a list of numeric vectors, where each vector is the vector of atoms (in ascending order) of the step function that is the CDF of the corresponding p-value.
method	correction method. See details.
cutoff	an upper limit for the p-values to be adjusted; set this (to non-default) if p-values above the cutoff may be viewed as corresponding to null hypotheses.
n	number of comparisons, must be at least length(p).

## Details

The adjustment methods include the step-up Benjamini & Hochberg (1995) procedure on mid P-values ("BHmidp"); the step-up procedure of Heyse (2011, "DBH"); the step-down Benjamini & Liu (1999) procedure on mid P-values ("BLmidp"); the step-down procedure of Heller & Gur (2011, "DBL"). For completeness, the step-up Benjamini & Hochberg (1995) procedure ("BH") and the step-down Benjamini & Liu (1999) procedure ("BL") are also provided.

For discrete tests, the procedures "BHmidP" and "BLmidP" have closer nominal FDR levels than "BH" and "BL" respectively. Moreover, when the p-values are independent procedure "DBL" has proven FDR control, along with procedures "BH" and "BL". For power comparisons across methods, see Heller & Gur (2011).

The cutoff can be set to a value between 0 and 1, usually 0.05 is a good conservative guess that will alleviate the computational burden without power loss. All unadjusted p-values above this value will not be adjusted, and will receive a default value of 1 in the output vector. The purpose of cutoff is to reduce substantially computational costs in very large number of tests.

n can be set to a value larger than length(p) which means the unobserved p-values are assumed to be equal to 1.

**Value**

A numeric vector of the adjusted p-values (of the same length as p).

**Note**

The function structure and code is mainly based on the code in [p.adjust](#) written by R Development Core Team. The BH method is identical to the code in [p.adjust](#).

**Author(s)**

Ruth Heller, Hadas Gur and Shay Yaacoby.

Maintainer: Shay Yaacoby <shay66@gmail.com>

**References**

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, **57**, 289–300.

Benjamini, Y., and Liu, W. (1999). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Statistical planning and inference*, **82**, 163–170.

Heller, R., and Gur, H. (2011). False discovery rate controlling procedures for discrete tests. *arXiv:1112.4627v1* [link](#).

Heyse, J. (2011). A false discovery rate procedure for categorical data. *Resent Advances in Biostatistics: False Discovery Rates, Survival Analysis, and Related Topics*, 43–58.

**See Also**

[p.adjust](#)

**Examples**

```
data(amesia)

A11 <- amnesia$AmnesiaCases
A21 <- sum(amesia$AllAdverseCases) - A11
A12 <- amnesia$AllAdverseCases - A11
A22 <- sum(amesia$AllAdverseCases) - sum(amesia$AmnesiaCases) - A12

## Entry j in each of the four vectors is the data for the test of no association
## between drug j and amnesia :
##
##      Drug j      Other Drugs
## Amnesia      A11[j]      A12[j]      A1.[j]
## Other Adverse events A21[j]      A22[j]      A2.[j]
##
##              n          N-n          N

## For example, the 2X2 contingency table to test the hypothesis of
## amnesia adverse drug reaction in the drug "ZOPICLONE":
matrix(c(A11[2444], A21[2444], A12[2444], A22[2444]),nrow = 2)

A1. <- sum(amesia$AmnesiaCases)
```

```

A2. <- sum(amesia$AllAdverseCases) - A1.
n <- A11 + A12
k <- pmin(n,A1.)

pCDFlist <- list()
pvec <- numeric(nrow(amesia))

## Calculation of the p-values and the p-values CDFs:

for (i in 1:nrow(amesia))
{
  x <- 0:k[i]
  pCDFlist[[i]] <- dhyper(x ,A1. , A2. ,n[i]) + phyper(x ,A1. ,A2. ,n[i] ,lower.tail = FALSE)
  pCDFlist[[i]] <- rev(pCDFlist[[i]])
  pvec[i] <- dhyper(A11[i] ,A1. ,A2. ,n[i]) + phyper(A11[i] ,A1. ,A2. ,n[i] ,lower.tail = FALSE)
}

pBH <- p.discrete.adjust(pvec, pCDFlist, method = "BH")
pBL <- p.discrete.adjust(pvec, pCDFlist, method = "BL")
pBHmidp <- p.discrete.adjust(pvec, pCDFlist, method = "BHmidp")
pBLmidp <- p.discrete.adjust(pvec, pCDFlist, method = "BLmidp")
pDBH <- p.discrete.adjust(pvec, pCDFlist, method = "DBH")
pDBL <- p.discrete.adjust(pvec, pCDFlist, method = "DBL")

## Number of rejected hypothesis at level 0.05:
q <- 0.05
sum(pBL <= q)    ## 16
sum(pBLmidp <= q) ## 17
sum(pDBL <= q)  ## 21
sum(pBH <= q)   ## 24
sum(pBHmidp <= q) ## 25
sum(pDBH <= q)  ## 27

## plotting:
o = order(pvec)

matplot(1:length(pvec), cbind(pvec[o], pBL[o], pBLmidp[o], pDBL[o], pBH[o], pBHmidp[o], pDBH[o]),
        type = "l", lty = c(4,3,3,3,2,2,2),
        col = c("#4735B2", "#B25A00", "#24B200", "#106B99", "#B25A00", "#24B200", "#106B99"),
        xlim = c(1,100), xlab = "Rank", ylab = "Adjusted p-values")
abline(0.05,0,col = "grey")
legend("bottomright", legend=c("pvec-unadjusted", "pBL", "pBLmidp", "pDBL", "pBH", "pBHmidp", "pDBH"),
      lty = c(4,3,3,3,2,2,2),
      col = c("#4735B2", "#B25A00", "#24B200", "#106B99", "#B25A00", "#24B200", "#106B99"))

```

# Index

## \*Topic **datasets**

amnesia, [2](#)

amnesia, [2](#)

as.numeric, [3](#)

discreteMTP (discreteMTP-package), [1](#)

discreteMTP-package, [1](#)

NA, [3](#)

p.adjust, [1](#), [3](#), [4](#)

p.discrete.adjust, [1](#), [3](#)