Package 'simpleFDR'

November 4, 2021

Type Package

Title Simple False Discovery Rate Calculation

Description Using the adjustment method from Benjamini & Hochberg (1995) <doi:10.1111/j.2517-6161.1995.tb02031.x>, this package determines which variables are significant under repeated testing with a given dataframe of p values and an user defined ``q'' threshold. It then returns the original dataframe along with a significance column where an asterisk denotes a significant p value after FDR calculation, and NA denotes all other p values. This package uses the Benjamini & Hochberg method specifically as described in Lee, S., & Lee, D. K. (2018) <doi:10.4097/kja.d.18.00242>.

Author Stephen C Wisser

Version 1.1

RoxygenNote 7.1.2

Imports dplyr,tidyr

Maintainer Stephen Wisser <swisser98@gmail.com>

License MIT + file LICENSE

Date 2021-11-01

NeedsCompilation no

Repository CRAN

Date/Publication 2021-11-04 07:20:02 UTC

R topics documented:

simFDR *Simple False Discovery Rate (FDR) Calculation*

Description

Using the Benjamini & Hochberg adjustment method, determine which variables are significant under repeated testing with a given dataframe of p values and an user defined "q" threshold.

Usage

```
simFDR(df, q = 0.05, sig_only = TRUE)
```

Arguments

df	dataframe with variable names in column 1, and p values in column 2. For dataframes with more than these 2 columns, the additional columns will be ignored. Example: df<- data.frame("variable"= c("a","b","c","d","e","f","g","h","i","j","k"),"p_value" = c(0.04,0.03,0.04,0.02,0.03,0.02,0.02,0.01,0.04,0.1,0.02))
q	user defined FDR threshold. Defaults to 0.05.
sig_only	logical value indicating whether to return just the variables that are significant, or all input variables. If TRUE, only significant variables are returned. If FALSE, all variables are returned with the significant variables at the top. Defaults to TRUE.

Value

Returns the original dataframe with a significance column where an asterisk denotes a significant p value after FDR calculation, and NA denotes all other p values.

Author(s)

Stephen C Wisser

References

Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test?. Korean journal of anesthesiology, 71(5), 353.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Examples

```
df <- data.frame("variable"= c("a","b","c","d","e"),"p_value" = c(0.04,0.03,0.04,0.02,0.03))

# defaults to q = 0.05 and shows only significant p values
FDR_values <- simFDR(df)

# q = 0.1 and shows only significant p values
FDR_values <- simFDR(df, q = 0.1)

# q = 0.05 and shows all p values, with significant p values at the top
FDR_values <- simFDR(df, sig_only = FALSE)
```

Index