Introduction to BonEV package

Dongmei Li

August 10, 2024

Clinical and Translational Science Institute, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642-0708

Contents

1 Overview

This document provides an introduction to the `BonEV` package. The `BonEv` package calculates the adjusted P-values from user-provided raw P-values through the Bon-EV multiple testing procedure that controls the false discovery rates at user-defined level alpha. The Bon-EV multiple testing procedure is developed based on the Bonferroni procedure with integrated estimates from the Benjamini-Hochberg procedure and the Storey's q-value procedure. It controls false discovery rates through controlling the expected number of false discoveries.

2 Getting started

The `BonEV` package can be installed and loaded through the following R code.
Install the `BonEV` package with:

```
> install.packages("BonEV")
```

Load the `BonEV` package with:

```
> library(BonEV)
```

3 `Bon_EV` function

There is one function in the `BonEV` package: `Bon_EV`. The function requires raw P-values in the vector format and user-defined alpha level for false discovery rates control. `Bon_EV` will generate adjusted

P-values from the Bon-EV multiple testing procedure that is developed based on the Bonferroni procedure with integrated estimates from the Benjamini-Hochberg procedure and the Storey's q-value procedure. `Bon_EV` controls false discovery rates through controlling the expected number of false discoveries.
The following is an example using the `Bon_EV` function. The raw P-values in the hedenfalk data set from the qvalue package are used as the input to get adjusted P-values from the Bon-EV multiple testing procedure with the false discovery rate controlled at level alpha = 0.05. Then, the adjusted P-values from the Bon-EV multiple testing procedure are compared with adjusted P-values obtained from the Benjamini-Hochberg and Storey's q-value procedures.

```
> library(qvalue)
> data(hedenfalk)
> library(BonEV)
> pvalues <- hedenfalk$p
> adjp <- Bon_EV(pvalues, 0.05)
> summary(adjp)

            Length Class  Mode
raw_P_value 3170   -none- numeric
BH_adjp     3170   -none- numeric
Storey_adjp 3170   -none- numeric
Bon_EV_adjp 3170   -none- numeric

> results <- cbind(adjp$raw_P_value, adjp$BH_adjp, adjp$Storey_adjp, adjp$Bon_EV_adjp)
> colnames(results) <- c("raw_P_value", "BH_adjp", "Storey_adjp", "Bon_EV_adjp")
> results[1:20,]

       raw_P_value    BH_adjp Storey_adjp Bon_EV_adjp
 [1,] 0.0121261830 0.13164384  0.08819163  0.27395698
 [2,] 0.0750252366 0.31252300  0.20936729  1.00000000
 [3,] 0.9949211356 0.99712298  0.66799864  1.00000000
 [4,] 0.0417854890 0.24127505  0.16163643  0.94402555
 [5,] 0.8458138801 0.94409507  0.63247386  1.00000000
 [6,] 0.2519242902 0.55487215  0.37172329  1.00000000
 [7,] 0.6586561514 0.85571311  0.57326449  1.00000000
 [8,] 0.0656813880 0.29288112  0.19620868  1.00000000
 [9,] 0.1232681388 0.40326109  0.27015510  1.00000000
[10,] 0.0007129338 0.03455882  0.02315186  0.01610673
[11,] 0.0883974763 0.34467564  0.23090718  1.00000000
[12,] 0.0073817035 0.10263158  0.06875557  0.16676882
[13,] 0.2710000000 0.56872932  0.38100657  1.00000000
[14,] 0.9749810726 0.99124118  0.66405827  1.00000000
[15,] 0.2497097792 0.55277933  0.37032126  1.00000000
[16,] 0.7734763407 0.91832210  0.61520787  1.00000000
[17,] 0.0361829653 0.22569472  0.15119876  0.81745229
[18,] 0.0017507886 0.05441176  0.03645186  0.03955414
```

```
[19,] 0.0884668770 0.34467564  0.23090718  1.00000000
[20,] 0.1380883281 0.42947059  0.28771352  1.00000000

> summary(results)

  raw_P_value           BH_adjp        Storey_adjp        Bon_EV_adjp
 Min.   :0.0000032   Min.   :0.0100   Min.   :0.006699   Min.   :0.0000713
 1st Qu.:0.0845647   1st Qu.:0.3379   1st Qu.:0.226394   1st Qu.:1.0000000
 Median :0.2998155   Median :0.5993   Median :0.401516   Median :1.0000000
 Mean   :0.3718702   Mean   :0.5764   Mean   :0.386171   Mean   :0.8831296
 3rd Qu.:0.6316112   3rd Qu.:0.8418   3rd Qu.:0.563918   3rd Qu.:1.0000000
 Max.   :0.9998517   Max.   :0.9999   Max.   :0.669827   Max.   :1.0000000

> ##Compare with Benjami-Hochberg and Storey's q-value procedures
> sum(adjp$raw_P_value <= 0.05)

[1] 606

> sum(adjp$BH_adjp <= 0.05)

[1] 94

> sum(adjp$Storey_adjp <= 0.05)

[1] 162

> sum(adjp$Bon_EV_adjp <= 0.05)

[1] 120
```