RUCStatBeamer—Typst Template

Make your slides with Typst

Song Wenxuan

Renmin University of China

2025-02-03

Outline

Outline

How to choose a threshold?

- Control Per-Comparison Type I Error (**PCER**)
  1. a.k.a. "*uncorrected testing*", many type I errors
  2. Gives $\mathbb{P}\{\mathrm{FD}_i > 0\} \leqslant \alpha$ marginally for all $1 \leqslant i \leqslant m$
- Control Familywise Type I Error (**FWER**)
  1. e.g. Bonferroni method, or using per-comparison significance level $\frac{\alpha}{m}$
  2. Guarantees $\mathbb{P}\{\mathrm{FD} > 0\} \leqslant \alpha$
- Control False Discovery Rate (**FDR**)
  1. First defined by Benjamini & Hochberg [1]
  2. Guarantees $\mathrm{FDR} \equiv \mathbb{E}\left(\frac{\mathrm{FD}}{D}\right) \leqslant \alpha$

BH Procedure

Theorem 1

The Benjamini–Hochberg procedure (BH step-up procedure) controls the FDR at level α for independent multiple tests.

Visualization and Algorithm

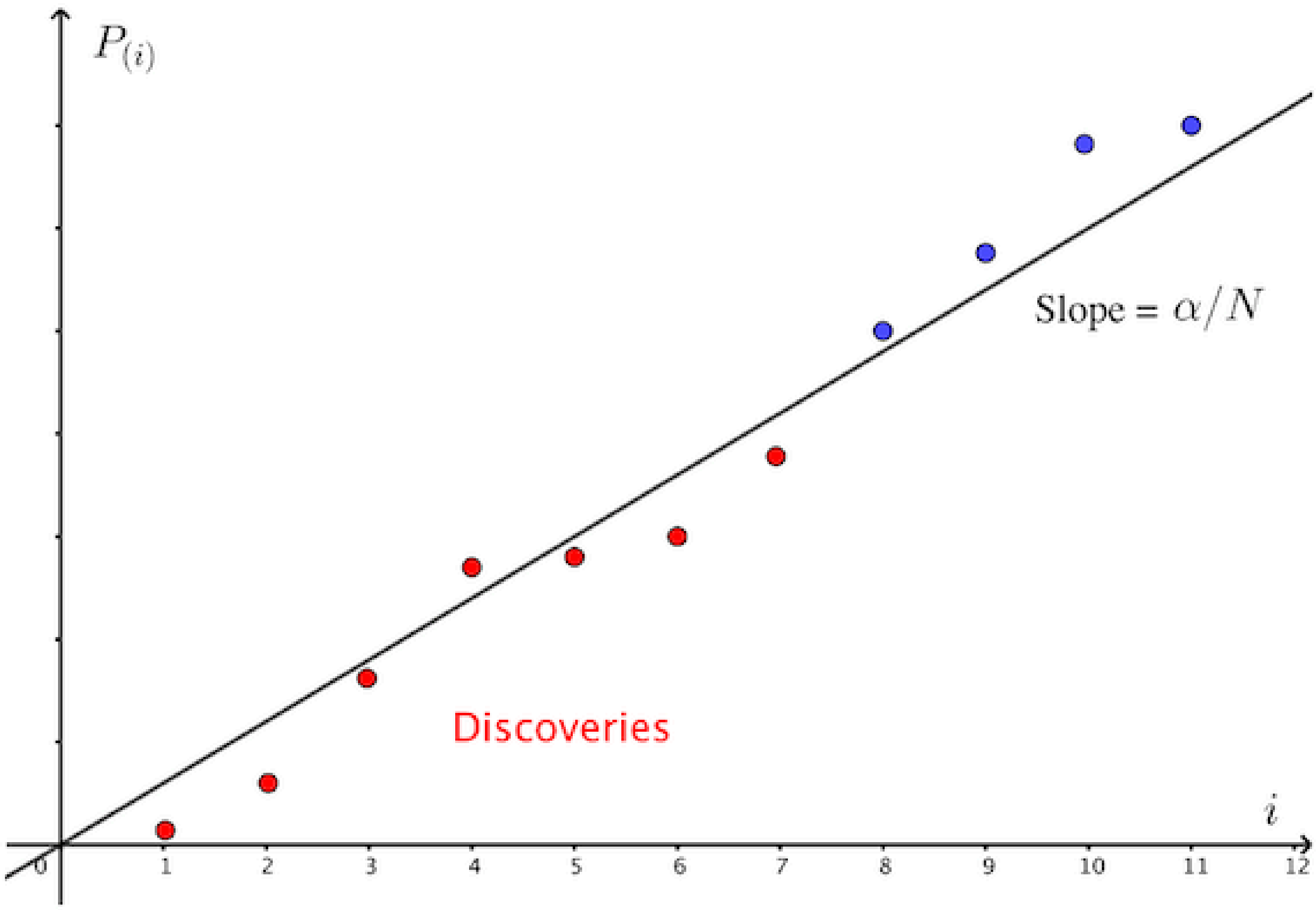

Figure 1: BH Procedure

Visualization and Algorithm

1. For a given α, find the largest k such that $P_{(k)} \leqslant \frac{\cancel{k}}{\cancel{m}} \frac{k}{m} \alpha$.
2. Reject the null hypothesis (i.e., declare discoveries) for all $H_{(i)}$ for $i = 1, \ldots, k$.

Outline

R Code

```
bh <- function() {
    UseMethod("bh")
}

bh.func  <-  function(pv,  alpha  =
0.05) {
    m <- length(pv)
    i <- 1:m
    sorted_pv <- sort(pv)
    if (sorted_pv[1] > alpha / m) {
        return(rep(0, m))
    }
     k  <-  max(i[sorted_pv  <=  i  /  m
* alpha])
    criterion <- sorted_pv[k]
    return(1 * (pv <= criterion))
}
```

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri.

$\boldsymbol{P}$ P-value, or realization of a P-variable.

$\boldsymbol{E}$ E-value, or realization of a E-variable.

$\boldsymbol{Q}$ Q-value, or realization of a Q-variable.

Subsection B.2

You can use `#pause` to pause display some proof. 🤔

and this completes the proof. □

Subsection B.2

You can use `#pause` to pause display some proof. 🤔

Proof: Let $\alpha_r = \alpha r / K$ for $r \in \mathcal{K}$. We can write

$$\mathbb{E}\left[\frac{F_{\mathcal{D}}}{R_{\mathcal{D}}}\right] = \mathbb{E}\left[\frac{\sum_{k \in \mathcal{N}} \mathbb{1}_{\{P_k \leqslant \alpha_{R_{\mathcal{D}}}\}}}{R_{\mathcal{D}}}\right] = \sum_{k \in \mathcal{N}} \sum_{r=1}^{K} \frac{1}{r} \mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_{\mathcal{D}} = r\}}\right]. \qquad (1)$$

and this completes the proof. □

Subsection B.2

You can use `#pause` to pause display some proof. 🤔

Proof: Let $\alpha_r = \alpha r / K$ for $r \in \mathcal{K}$. We can write

$$\mathbb{E}\left[\frac{F_{\mathcal{D}}}{R_{\mathcal{D}}}\right] = \mathbb{E}\left[\frac{\sum_{k \in \mathcal{N}} \mathbb{1}_{\{P_k \leqslant \alpha_{R_{\mathcal{D}}}\}}}{R_{\mathcal{D}}}\right] = \sum_{k \in \mathcal{N}} \sum_{r=1}^{K} \frac{1}{r} \mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_{\mathcal{D}} = r\}}\right]. \quad (1)$$

For $k \in \mathcal{N}$, let R_k be the number of rejection from the BH procedure if it is applied to $\mathbf{P}$ with P_k replaced by 0. Note that $\{P_k \leqslant \alpha_r, R_{\mathcal{D}} = r\} = \{P_k \leqslant \alpha_r, R_k = r\}$ for each k, r. Hence, we have

$$\mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_{\mathcal{D}} = r\}}\right] = \mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_k = r\}}\right]. \quad (2)$$

and this completes the proof. □

Subsection B.2

You can use `#pause` to pause display some proof. 🤔

Proof: Let $\alpha_r = \alpha r / K$ for $r \in \mathcal{K}$. We can write

$$\mathbb{E}\left[\frac{F_{\mathcal{D}}}{R_{\mathcal{D}}}\right] = \mathbb{E}\left[\frac{\sum_{k \in \mathcal{N}} \mathbb{1}_{\{P_k \leqslant \alpha_{R_{\mathcal{D}}}\}}}{R_{\mathcal{D}}}\right] = \sum_{k \in \mathcal{N}} \sum_{r=1}^{K} \frac{1}{r} \mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_{\mathcal{D}} = r\}}\right]. \quad (1)$$

For $k \in \mathcal{N}$, let R_k be the number of rejection from the BH procedure if it is applied to $\mathbf{P}$ with P_k replaced by 0. Note that $\{P_k \leqslant \alpha_r, R_{\mathcal{D}} = r\} = \{P_k \leqslant \alpha_r, R_k = r\}$ for each k, r. Hence, we have

$$\mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_{\mathcal{D}} = r\}}\right] = \mathbb{E}\left[\mathbb{1}_{\{P_k \leqslant \alpha_r\}} \mathbb{1}_{\{R_k = r\}}\right]. \quad (2)$$

Putting this into Equation 1, we get

$$\mathbb{E}\left[\frac{F_{\mathcal{D}}}{R_{\mathcal{D}}}\right] = \frac{\alpha}{K} \sum_{k \in \mathcal{N}} \sum_{r=1}^{K} \mathbb{P}(R_k = r) = \frac{K_0 \alpha}{K},$$

and this completes the proof. □

Outline

Appendix

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

[2] Bonferroni C. Teoria statistica delle classi e calcolo delle probabilita. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze. 1936;8:3–62.

[3] Dwork C, McSherry F, Nissim K, et al. Calibrating noise to sensitivity in private data analysis. In Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006 Proceedings 3. 2006. pp. 265–284.

[4] Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979;6(2):65–70.

[5] Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika. 1988;75(4):800–802.