Multiple Testing

1 Problem Formulation

Suppose there are n hypotheses $H_{0,\ell}$. Types of outcomes in multiple testing can be summarized as follows:

	H_0 accepted	H_0 rejected	Total
H_0 true	U	V	n_0
H_0 false	T	S	$n - n_0$
Total	$n - R$	R	n

- U, V, T, S are unobserved random variables;
- R is an observed random variable.

2 Familywise Error Rate (FWER)

The familywise error rate is defined as

$$FWER = \mathbb{P}(V \geq 1),$$

i.e., the probability of at least one false rejection. Classical multiple testing methods aim to control FWER in a strong sense, i.e., under all configurations of the true and false hypotheses being tested.

2.1 Bonferroni correction

In Bonferroni correction, $H_{0,i}$ is rejected iff $p_i < \alpha/n$.

Theorem 2.1 *Bonferroni's method controls FWER at level α.*

Proof Let I be the set of true nulls. Let $|I| = n_0$. For each $i \in I$, let V_i be the event that $H_{0,i}$ is rejected, i.e., $p_i < \alpha/n$. Then

$$\mathbb{P}(V \geq 1) = \mathbb{P}\left(\bigcup_{i \in I} V_i\right) \leq \sum_{i \in I} \mathbb{P}_{0,i}(V_i) \leq \sum_{i \in I} \frac{\alpha}{n} = n_0 \frac{\alpha}{n} \leq \alpha.$$

■

2.2 Holm's procedure

Let the order statistics of the n p-values be

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)},$$

with corresponding hypotheses

$$H_{0,(1)}, H_{0,(2)}, \ldots, H_{0,(n)}.$$

- **Step 1**: If $p_{(1)} \leq \frac{\alpha}{n}$, reject $H_{0,(1)}$ and go to Step 2; otherwise, accept $H_{0,(1)}$,..., $H_{0,(n)}$ and stop;
- ....
- **Step i**: If $p_{(i)} \leq \frac{\alpha}{n-i+1}$, reject $H_{0,(i)}$ and go to the next step; otherwise, accept $H_{0,(i)}$,..., $H_{0,(n)}$ and stop;
- ...
- **Step n**: If $p_{(n)} \leq \alpha$, reject $H_{0,(n)}$; otherwise, accept $H_{0,(n)}$ and stop.

Not as conservative as Bonferroni: Reject more and more powerful.

Theorem 2.2 *Holm's procedure controls FWER strongly.*

Proof Let I be the set of true nulls. Let $|I| = n_0$. Denote $i_0 = \arg\min_{i\in I} p_{(i)}$. Then $p_{(i_0)}$ is the minimum p-value among true nulls, and obviously $i \leq n - n_0 + 1 \Longrightarrow n_0 \leq n - i_0 + 1$.

Holm's procedure commits a false rejection only if

$$p_{(1)} \leq \frac{\alpha}{n}, p_{(2)} \leq \frac{\alpha}{n-1}, \ldots, p_{(i_0)} \leq \frac{\alpha}{n-i_0+1}.$$

Then

$$p_{(i_0)} \leq \frac{\alpha}{n-i_0+1} \leq \frac{\alpha}{n_0}.$$

Then, the probability of a false rejection is bounded above by

$$\mathbb{P}\left[\min_{i\in I} p_i \leq \frac{\alpha}{n_0}\right] = \mathbb{P}\left[\bigcup_{i\in I}\left\{p_i \leq \frac{\alpha}{n_0}\right\}\right] \leq \sum_{i\in I}\mathbb{P}(p_i \leq \frac{\alpha}{n_0}) \leq \sum_{i\in I}\frac{\alpha}{n_0} = \alpha.$$

■

3 False Discovery Rate

When the number of tests are in tens of thousands or higher, control of FWER is so stringent that individual departures from null have little chance of being detected. Benjamini and Hochberg 1995 proposed a new point of view: control expected proportion of errors among rejected hypotheses.

- $FWER = \mathbb{P}(V \geq 1)$
- $FDP = \frac{V}{\max(R,1)} = \begin{cases} \frac{V}{R} & \text{if } R \geq 1 \\ 0 & \text{if } R = 1, \end{cases}$ which is an unobserved random variable.
- $FDR = \mathbb{E}[FDP] = \mathbb{E}\left[\frac{V}{\max(R,1)}\right]$.

There are two properties:

- If all hypotheses are null, then FDR is equivalent to FWER. In fact, given the facts $V = R$ and

$$\begin{cases} V = 0 & \text{iff} \quad FDP = 0 \\ V \geq 1 & \text{iff} \quad FDP = 1. \end{cases}$$

  Then,

$$\mathbb{P}(V \geq 1) = \mathbb{E}[FDP] = FDR.$$

- When $n_0 < n$, we have
$$\begin{cases} V = 0 & \text{iff} \quad FDP = 0 \\ V \geq 1 & \text{iff} \quad FDP \leq 1. \end{cases}$$
Then $1_{\{V \geq 1\}} \geq FDP$, which implies that
$$FWER = \mathbb{P}(V \geq 1) = \mathbb{E}[1_{V \geq 1}] \geq \mathbb{E}[FDP] = FDR.$$

Therefore, any procedure controlling FDR can be less stringent and more powerful.

3.1 Benjamini Hochberg Procedure

Consider the ordered P-values
$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}.$$
Fix $q \in [0, 1]$. Let i_0 be the largest index for which $p_{(i)} \leq \frac{i}{n} q$, and then we reject all $H_{0,(i)}$ with $i \leq i_0$.

Theorem 3.1 *For independent test statistics (p-values), the $BH(q)$ procedure controls FDR at level q in that*
$$FDR \leq \frac{n_0}{n} q \leq q.$$

This holds for all configurations of false hypotheses. BH procedure is Sines procedure for weak FWER control [Sines 86].
Proof WLOG, let $I = \{1, \ldots, n_0\}$. For any integer r,
$$R = r \iff p_{(r)} \leq q\frac{r}{n}, \text{ and } p_{(s)} \geq q\frac{s}{n}, \quad \forall s > r.$$

Therefore,
$$\begin{aligned} \left\{p_1 \leq \frac{qr}{n},\ R = r\right\} &= \left\{p_1 \leq \frac{qr}{n}, \quad p_{(r)} \leq \frac{rq}{n}, \text{ and } p_{(s)} > \frac{qs}{n}, \quad \forall s > r\right\} \\ &= \left\{p_1 \leq \frac{qr}{n}, \quad p^{(1)}_{(r-1)} \leq \frac{qr}{n}, \text{ and } p^{(1)}_{(s)} > \frac{q(s+1)}{n}, \quad \forall s > r-1\right\}, \end{aligned}$$

where $p^{(1)}_{(1)}, \ldots, p^{(1)}_{(n-1)}$ are the order statistics excluding p_1. Denote
$$C^{(1)}_r = \left\{p^{(1)}_{(1)}, \ldots, p^{(1)}_{(r-1)} \leq \frac{qr}{n},\ p^{(1)}_{(r)} > \frac{q(r+1)}{n}, \ldots, p^{(1)}_{(n-1)} > q\right\}.$$

Then $C^{(1)}_1$, $C^{(1)}_2$, ... $C^{(1)}_n$ are a partition of the whole sample space Ω. Moreover, we have
$$\left\{p_1 \leq \frac{qr}{n},\ R = r\right\} = \left\{p_1 \leq \frac{qr}{n}\right\} \cap C^{(1)}_r.$$

Then,

$$
\begin{aligned}
FDR &= \sum_{r=1}^{n} \mathbb{E}\left\{\frac{V}{r} 1_{\{R=r\}}\right\} \\
&= \sum_{r=1}^{n} \frac{1}{r}\, \mathbb{E}\left[\sum_{i=1}^{n_0} 1_{\{p_i \le \frac{qr}{n}\}} 1_{\{R=r\}}\right] \\
&= \sum_{r=1}^{n} \frac{n_0}{r}\, \mathbb{E}\left[1_{\{p_1 \le \frac{qr}{n}\}} 1_{\{R=r\}}\right] \\
&= \sum_{r=1}^{n} \frac{n_0}{r}\, \mathbb{P}\left(\left\{p_1 \le \frac{qr}{n}, R=r\right\}\right) \\
&= \sum_{r=1}^{n} \frac{n_0}{r}\, \mathbb{P}\left(\left\{p_1 \le \frac{qr}{n}\right\} \cap C_r^{(1)}\right) \\
&= \sum_{r=1}^{n} \frac{n_0}{r}\, \mathbb{P}\left(\left\{p_1 \le \frac{qr}{n}\right\}\right) \mathbb{P}\left(C_r^{(1)}\right) \\
&= \sum_{r=1}^{n} \frac{n_0}{r} \frac{qr}{n}\, \mathbb{P}\left(C_r^{(1)}\right) \\
&= \frac{qn_0}{n} \sum_{r=1}^{n} \mathbb{P}\left(C_r^{(1)}\right) = \frac{qn_0}{n}.
\end{aligned}
$$

■

3.2 Controlling of FDR under PRDS

Without independence, we still have the expression of FDR by BH procedure

$$
\begin{aligned}
FDR &= \sum_{r=1}^{n} \frac{1}{r} \sum_{i=1}^{n_0} \mathbb{P}\left(\left\{p_i \le \frac{qr}{n}, R=r\right\}\right) \\
&= \sum_{r=1}^{n} \frac{1}{r} \sum_{i=1}^{n_0} \mathbb{P}\left(\left\{p_i \le \frac{qr}{n}\right\} \cap C_r^{(i)}\right) \\
&= \sum_{i=1}^{n_0} \sum_{r=1}^{n} \frac{1}{r}\, \mathbb{P}\left(\left\{p_i \le \frac{qr}{n}\right\} \cap C_r^{(i)}\right),
\end{aligned}
$$

where

$$
C_r^{(i)} = \left\{p_{(1)}^{(i)}, \ldots, p_{(r-1)}^{(i)} \le \frac{qr}{n},\ p_{(r)}^{(i)} > \frac{q(r+1)}{n}, \ldots, p_{(n-1)}^{(i)} > q\right\},
$$

and $p_{(1)}^{(i)}, \ldots, p_{(n-1)}^{(i)}$ are the order p-values excluding p_i.

In order to control FDR, let's define the concept of PRDS. First, for two jointly distributed random variables X and Y, we say X is PRDS on Y, if for any α and $y \le y'$, there holds

$$
\mathbb{P}(X \ge \alpha | Y \le y) \le \mathbb{P}(X \ge \alpha | Y \le y').
$$

That is, X gets stochastically larger when Y increases.

In the case that $\vec{X}$ is a random vector, in order to introduce the concept of PRDS, we first define the concept of increasing sets: A set $D \in \mathbb{R}^n$ is said to be an increasing set, if

$$\vec{x} \in D \text{ and } \vec{y} \geq \vec{x} \implies \vec{y} \in D.$$

Here we assume the p-values $p_1, \ldots p_n$ satisfy the PRDS on $I_0 = \{1, \ldots, n_0\}$, i.e., for any $i = 1, \ldots, n_0$, and any increasing set $D \in \mathbb{R}^n$,

$$\mathbb{P}(\vec{p} \in D | p_i \leq x)$$

is increasing in x. Here $\vec{p} = [p_1, \ldots, p_n]^\top$.

For each $i = 1, \ldots, n_0$, we know that $C_1^{(i)}$, $C_2^{(i)}$, ... $C_n^{(i)}$ are a partition of the whole sample space. Moreover, we have $C_1^{(i)}$, $C_1^{(i)} \cup C_2^{(i)}$, $C_1^{(i)} \cup C_2^{(i)} \cup C_3^{(i)}$, ... are all increasing sets. Notice that

$$\begin{aligned}
FDR &= \sum_{i=1}^{n_0} \sum_{r=1}^{n} \frac{1}{r} \mathbb{P}\left(\left\{p_i \leq \frac{qr}{n}\right\} \cap C_r^{(i)}\right) \\
&= \sum_{i=1}^{n_0} \sum_{r=1}^{n} \frac{1}{r} \mathbb{P}\left(p_i \leq \frac{qr}{n}\right) \frac{\mathbb{P}\left(\left\{p_i \leq \frac{qr}{n}\right\} \cap C_r^{(i)}\right)}{\mathbb{P}\left(p_i \leq \frac{qr}{n}\right)} \\
&\leq \sum_{i=1}^{n_0} \sum_{r=1}^{n} \frac{1}{r} \frac{qr}{n} \mathbb{P}\left(C_r^{(i)} \middle| p_i \leq \frac{qr}{n}\right) \\
&= \frac{q}{n} \sum_{i=1}^{n_0} \sum_{r=1}^{n} \mathbb{P}\left(C_r^{(i)} \middle| p_i \leq \frac{qr}{n}\right).
\end{aligned}$$

By the PRDS assumption, we have

$$\begin{aligned}
&\sum_{r=1}^{n} \mathbb{P}\left(C_r^{(i)} \middle| p_i \leq \frac{qr}{n}\right) \\
&= \mathbb{P}\left(C_1^{(i)} \middle| p_i \leq \frac{q}{n}\right) + \mathbb{P}\left(C_2^{(i)} \middle| p_i \leq \frac{2q}{n}\right) + \mathbb{P}\left(C_3^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \ldots + \mathbb{P}\left(C_n^{(i)} | p_i \leq q\right) \\
&\leq \mathbb{P}\left(C_1^{(i)} \middle| p_i \leq \frac{2q}{n}\right) + \mathbb{P}\left(C_2^{(i)} \middle| p_i \leq \frac{2q}{n}\right) + \mathbb{P}\left(C_3^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \ldots + \mathbb{P}\left(C_n^{(i)} | p_i \leq q\right) \\
&= \mathbb{P}\left(C_1^{(i)} \cup C_2^{(i)} \middle| p_i \leq \frac{2q}{n}\right) + \mathbb{P}\left(C_3^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \ldots + \mathbb{P}\left(C_n^{(i)} | p_i \leq q\right) \\
&\leq \mathbb{P}\left(C_1^{(i)} \cup C_2^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \mathbb{P}\left(C_3^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \ldots + \mathbb{P}\left(C_n^{(i)} | p_i \leq q\right) \\
&= \mathbb{P}\left(C_1^{(i)} \cup C_2^{(i)} \cup C_3^{(i)} \middle| p_i \leq \frac{3q}{n}\right) + \ldots + \mathbb{P}\left(C_n^{(i)} | p_i \leq q\right) \\
&\leq \ldots \\
&\leq \mathbb{P}\left(C_1^{(i)} \cup C_2^{(i)} \cup C_3^{(i)} \cup \ldots \cup C_n^{(i)} | p_i \leq q\right) = 1.
\end{aligned}$$

Then we have

$$FDR \leq \frac{n_0 q}{n}.$$