STAT-511 **2018 Spring**

Lecture 5: Multiple Tests

Instructor: Naveen N. Narisetty Scribe: Yutong Dai Last Modified: 2018-05-06

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, and they may contain factual and/or typographic errors.*

Suppose have $H_{0i}, i = 1, \dots, n$ hypotheses, and there are in total n_0 true null hypotheses, but n_0 cannot be observed. Then there are only 4 possible outcomes for each H_{0i}:

	fail to reject H_{0i}	H_{0i} is rejected	total
H_{0i} is true		V	n_0
H_{0i} is false	U		n-n_0
total	n-R	R	n

We want to test the intersection null $H_0 : \cap_{i=1}^{n} H_{0i}$ and assume the p-values $\{p_i\}_{i=1}^{n}$ are given.

5.1 Control the Family-wise Error Rate (FWER)

Definition 5.1 FWER is defined as the probability of making any false rejection. Without loss of generality, assume $H_{0i}, i = 1, \dots, n_0$ are true null hypotheses, then

$$\text{FWER} = P(V > 0)$$

5.1.1 Bonferroni Procedure

To control FWER effectively, one of the simplest methods is *Bonferroni's global test*, or *Bonferroni Procedure*. Given a significant level α, it rejects H_0 when there exists any p_i s.t. $p_i < \frac{\alpha}{n}$.

Theorem 5.2 Bonferroni Procedure always controls FWER at level α.

Proof: Without loss of generality, assume $H_{0i}, i = 1, \dots, n_0$ are true null hypotheses.

$$\begin{aligned}
\text{FWER} = P(V > 0) =& P(\cup_{i=1}^{n_0}[p_i \leq \frac{\alpha}{n}]) \\
\leq& \sum_{i=1}^{n_0} P(p_i \leq \frac{\alpha}{n}) \\
\overset{(1)}{=}& \frac{n_0}{n}\alpha \leq \alpha,
\end{aligned}$$

where (1) holds as null hypotheses $H_{0i}, i = 1, \dots, n_0$ are true, so $P(p_i < \frac{\alpha}{n}) = \frac{\alpha}{n}$. ■

5.1.2 Fisher's Combination Test

Another way to do a global testing is *Fisher's Combination Test.* Although there's no theorem to ensure it controls FWER, it does complement Bonferroni's Test in some cases, where there are many small effects that are not significant enough to be rejected by Bonferroni Procedure.

Lemma 5.3 Let $X_i \overset{iid}{\sim} Unif(0,1)$, then $-2\sum_{i=1}^{n} \log X_i \sim \chi^2_{2n}$.

So if all H_{0i}'s are independent, theb under H_0(all H_{0i}'s are true),

$$-2\sum_{i=1}^{n} \log p_i \sim \chi^2_{2n}.$$

Therefore, given a significant level α, Fisher's Test rejects H_0 when $-2\sum_{i=1}^{n} \log p_i > \chi^2_{2n}(1-\alpha)$.

When comparing these two global testing procedures, we observe that Bonferroni's test only depends on the smallest p-value, i.e. a single strong effect. While Fisher's combination test aggregates all the p-values (in log-scale). It is more powerful than Bonferroni's if the alternative has many weak effects (i.e. small deviations from the null).

5.2 Control the False Discovery Rate (FDR)

The **False Discovery Rate (FDR)** is a method of conceptualizing the rate of *type I errors* in null hypothesis testing when conducting multiple testing. FDR-controlling procedures are designed to control the expected proportion of "discoveries" (rejected null hypotheses) that are false (incorrect rejections).

Definition 5.4 The FDR is defined as the expectation of False Discovery Proportion (FDP), $\frac{V}{R}$,

$$\text{FDR} = E[\frac{V}{R}] = E[\frac{\#\{TrueRejected\}}{\#\{Rejected\}}],$$

where the expectation is taken with respect to data, i.e., $x_i \sim$ some distributions. Specifically, FDR is defined to be 0 when R is zero.

Compared with FWER, FDR is more practical because it does not require making no mistakes. Instead, it only controls the proportion of mistakes.

Lemma 5.5 $\frac{1}{n}$ FWER $\leq$ FDR $\leq$ FWER

Proof: Note that

$$FWER = E[I(V>0)] \quad FDR = E[FDP \times I(V>0)],$$

and $\frac{V}{R} \geq \frac{1}{n}$ Therefore, $\frac{1}{n}$ FWER $\leq$ FDR $\leq$ FWER. ∎

5.2.1 Benjamini-Hochberg Procedure

Benjamini-Hochberg Procedure is formally described below.

Let $p_i, i = 1, \ldots, n$ be the p-values of H_{0i}.

- **Step1.** Sort $p_i, i = 1, \ldots, n$ so that $p_{(1)} < \cdots < p_{(n)}$.
- **Step2.** Find the largest i_1 such that $p_{(i_1)} < \frac{i_1}{n}\alpha$.
- **Step3.** Reject all H_{0i}'s that $p_i \leq p_{(i_1)}$.

Theorem 5.6 If p_i's are independent, then Benjamini-Hochberg Procedure controls FDR at level α.

Proof: Suppose there are n_0 true null hypotheses rejected. Then,

$$FDP = \frac{\#\{\text{True Rejected}\}}{\#\{\text{Rejected}\}} = \frac{\sum_{i:H_{0i}true} I(p_i \leq t)}{\sum_{i=1}^{n} I(p_i \leq t)}$$

Take expectaion on both sides,we derive

$$\text{FDR} \approx \frac{n_0 t}{\sum_{i=1}^{n} I(p_i \leq t)} \leq \frac{nt}{\sum_{i=1}^{n} I(p_i \leq t)}.$$

Take $t = p_{(j)}$, then we have $FDP \leq n\frac{p_{(j)}}{j}$. If we select the largest p-value so that $k\frac{p_{(j)}}{j} \leq \alpha$, then we can control FDP. ■

Remark: If we can know more about n_0, e.g. $n_0 \leq \frac{n}{2}$, then we can choose the largest $p_{(i_1)}$ such that $p_{(i_1)} \leq \frac{2i_1}{n}\alpha$.