Statistics 210A Lecture 27 Notes

Daniel Raban

December 2, 2021

1 Introduction to Multiple Hypothesis Testing

1.1 Correcting p-values to account for multiple hypotheses

Suppose $X \sim P_\theta \in \mathcal{P}$. We have hypotheses $H_{0,i} : \theta_{\in \theta_{0,i}}$ for $i = 1, \dots, m$. We will let

$$\mathcal{R}(X) = \{i : H_{0,i} \text{ is rejected}\}, \qquad \mathcal{H}_0(\theta) = \{i : \theta \in \Theta_{0,i}\}$$

and denote $R(X) = |\mathcal{R}|$ and $m_0 = |\mathcal{H}_0|$. The central issue is that the more hypotheses we test, the more likely we are to reject a hypothesis by chance.

Example 1.1. Let $X_i \stackrel{\text{iid}}{\sim} N(\theta_i, 1)$ with $H_{0,i} : \theta_i = 0$. Reject $H_{0,i}$ if $|X_i| > z_{\alpha/2}$. Then

$$\mathbb{P}_0(\text{any } H_{0,i} \text{ rejected}) = 1 - (1-\alpha)^m \xrightarrow{m \to \infty} 1.$$

Definition 1.1. The **familywise error rate (FWER)** is

$$\mathbb{P}_\theta(\text{any false rejections}) = \mathbb{P}_\theta(\mathcal{R} \cap \mathcal{H}_0 \neq \varnothing).$$

The classical view of multiple testing to say that we want

$$\sup_\theta \mathrm{FWER}_\theta \leq \alpha.$$

Remark 1.1. Why should we make a correction for the FWER? If you conduct 10 experiments and submit your analysis to a journal, they'll require you to make a familywise error correction. But if you submit the 1 experiment each to 10 journals, then no one will hassle you.

Sometimes, when we are testing many hypotheses, we care individually about each one. But sometimes, such as if we are testing hypotheses for a large number of genes, where we should expect most of our null hypotheses to be true, we might be okay with some percentage of our hypotheses being falsely rejected.

One way we can account for multiple hypotheses is to alter our p-values. Denote the p-values by $p_1(X), p_2(X), \dots, p_m(X)$. Here are some procedures for altering the p-values:

Example 1.2 (Sidák's correction)**.** Assume $p_i \geq U[0,1]$ for $i \in \mathcal{H}_0$. If the p_i are independent and we reject if $p_i \leq \widetilde{\alpha}_m$,

$$\begin{aligned} 1 - \alpha &= \mathbb{P}\theta(\text{no false rejection}) \\ &= \mathbb{P}(p_i \geq \widetilde{\alpha}_m \; \forall i \in \mathcal{H}_0) \\ &\geq (1 - \widetilde{\alpha}_m)^{m_0} \\ &\geq (1 - \widetilde{\alpha}_m)^{m}. \end{aligned}$$

If we solve this, we get

$$\widetilde{\alpha}_m = 1 - (1 - \alpha)^{1/m}.$$

If α is small, this is close to α/m.

Here is what we can do if we don't necessarily have independence.

Example 1.3 (Bonferroni correction)**.** Bonferroni rejects if $p_i \leq \alpha/m$. Then

$$\begin{aligned} \mathbb{P}_\theta(\text{any false rejection}) &= \mathbb{P}_\theta\left(\bigcup_{i \in \mathcal{H}_0} \{p_i \leq \alpha/m\}\right) \\ &\leq \sum_{i \in \mathcal{H}_0} \mathbb{P}_\theta(p_i \leq \alpha/m) \\ &\leq m_0 \cdot \frac{\alpha}{m} \\ &= \alpha \frac{m_0}{m}. \end{aligned}$$

The Bonferroni correction is still conservative. Here is a strictly better procedure:

Example 1.4 (Holm's procedure)**.**

Step 0: Order the p-values from small to large:

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}.$$

Let $H_{(i)}$ denote the hypothesis corresponding to $p_{(i)}$.

Step 1: If $p_{(i)} \leq \alpha/m$, reject $H_{(i)}$ and continue. Otherwise, stop and accept all null hypotheses.

$\vdots$

Step k: If $p_{(k)} \leq \frac{\alpha}{m-k+1}$, reject $H_{(k)}$ and continuous. Otherwise, stop and accept all hypotheses.

$\vdots$

Step m: If $p_{(m)} \leq \alpha$, reject $H_{(m)}$.

We can analyze this procedure by

$$R^{\text{Holm}} = \max\left\{r : p_{(i)} \leq \frac{\alpha}{m+i-1}\ \forall i \leq r\right\}.$$

We reject $H_{(1)}, \ldots, H_{(R^{\text{Holm}})}$.

Proposition 1.1. *Holm's procedure controls* FWER $\leq \alpha$.

Proof. Let $p_0^* = \min\{p_i : i \in \mathcal{H}_0\}$. Then

$$\mathbb{P}(p_0^* \leq \alpha/m_0) \leq \alpha$$

by the union bound. We claim that if $p_0^* > \alpha/m_0$, there are no false rejections. Let $k = \#\{i : p_i \leq p_0^*\} \leq m - m_0 + 1$. Then

$$p_{(k)} = p_0^* > \frac{\alpha}{m_0} \geq \frac{\alpha}{m-k+1}.$$

Then Holm makes $< k$ rejections. $\square$

1.2 The closure principle

Holm's procedure is an instance of the more general **closure principle**, which is used in a lot of modern developments in multiple testing methodology.

For $S \subseteq [m]$, let H_s be the hypothesis where all H_i are true for $i \in S$: $\theta \in \bigcap_{i\in S} \Theta_{0,i}$. Assume we have a level-α test for each subset. For example, we could reject H_s if $\min_{i\in S} p_i \leq \alpha/|S|$.

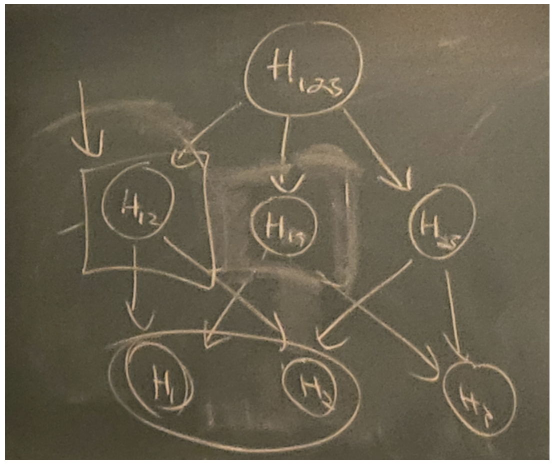
Step 1: Provisionally reject H_S if the corresponding marginal test ϕ_S rejects.

Step 2: Reject H_i if H_s is rejected for every $S \ni i$.

We can analyze the closure principle as follows:

$$\begin{aligned}\mathbb{P}(\text{any false rejections}) &\leq \mathbb{P}(H_{\mathcal{H}_0} \text{ is rejected in Step 1}) \\ &\leq \alpha.\end{aligned}$$

Here is a picture:



If H_1 and H_2 are the null hypotheses that are true, then they are protected as long as we don't reject the hypothesis $H_{1,2}$.

Remark 1.2. This might seem computationally inefficient, but as in our description of Holm's procedure, there can be computationally tractable ways to implementing this.

1.3 Testing with dependence

Example 1.5 (Scheffe's S-method)**.** Let $X \sim N_d(\theta, I_d)$ with $\theta \in \mathbb{R}^d$, and test $H_\lambda : \theta^\top \lambda = 0$ for $\lambda \in \mathbb{S}^{d-1}$ (this is uncountably infinitely many hypotheses). Reject H_λ if $(X^\top \lambda)^2 \geq \chi_d^2(\alpha) \approx 3 + 3\sqrt{d}$ if $\alpha = 0.05$. This controls the FWER because

$$\begin{aligned} \sup_{\lambda : \theta^\top \lambda = 0} (X^\top \lambda)^2 &\leq \sup_{\lambda \in \mathbb{S}^{d-1}} ((X - \theta)^\top \lambda)^2 \\ &= \|X - \theta\|^2 \\ &\sim \chi_d^2. \end{aligned}$$

Scheffe's method is a deduction from a spherical confidence region for θ: We can use $\|X - \theta\|^2 \sim \chi_d^2$ to get a confidence region for X. We get a confidence interval for θ^λ via

$$X^\top \lambda \pm \sqrt{\chi^2_d(\alpha)} \approx X^\top \lambda \pm \sqrt{d}.$$