Multiple testing issues

Motivation: differential gene expression

Suppose a biologist is interested in identifying genes which are *differentially expressed* under different biological treatments. The biologist observes 10 subjects under treatment A, and 10 subjects under treatment B. Gene expression measurements $X_{i,j}$ (treatment A) and $Y_{i,j}$ (treatment B) are recorded for 1000 different genes ($i = 1, \ldots, 1000, j = 1, \ldots, 10$).

For each gene i, the biologist tests $H_0 : \mu_{i,A} = \mu_{i,B}$, rejecting when the p-value is below a threshold α.

> If H_0 is actually true for all 1000 genes, how many false positives do we expect?

Motivation: multiple testing

> In what other settings might we test many hypotheses?

Family-wise error rate

Definition: Suppose we test m null hypotheses $H_{0,1}, \ldots, H_{0,m}$. The *family-wise error rate* is the probability of making *at least one* type I error:

$$FWER = P\left(\bigcup_{i=1}^{m} \{\text{reject } H_{0,i} | H_{0,i} \text{ is true}\}\right)$$

> Suppose all m tests are independent. For each test, we reject if the corresponding p-value $p_i < \alpha$. What is the FWER?

The Sidak correction

Suppose we test m null hypotheses $H_{0,1}, \ldots, H_{0,m}$. The *family-wise error rate* is the probability of making *at least one* type I error:

$$FWER = P\left(\bigcup_{i=1}^{m} \{\text{reject } H_{0,i} | H_{0,i} \text{ is true}\}\right)$$

> If all m hypotheses are independent, at what threshold α^* should we reject each test, such FWER = α?

The Bonferroni correction

Suppose we test m null hypotheses $H_{0,1}, \ldots, H_{0,m}$. The *family-wise error rate* is the probability of making *at least one* type I error:

$$FWER = P\left(\bigcup_{i=1}^{m} \{\text{reject } H_{0,i} | H_{0,i} \text{ is true}\}\right)$$

Holm's procedure

Suppose we test 5 hypotheses, and observe p-values 0.4, 0.01, 0, 0, 0. Does it still seem reasonable to use the Bonferroni cutoff $\alpha/5$ for each test?

Holm's procedure

Suppose we test m null hypotheses $H_{0,1}, \ldots, H_{0,m}$. Let p_i be the corresponding p-value for test i.

- Order the p-values $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$
- Let $i^* = \min \left\{ i : p_{(i)} > \frac{\alpha}{m-i+1} \right\}$
- Reject $H_{0,(i)}$ for all $i < i^*$

Claim: Holm's procedure controls FWER at level α

Class activity

https://sta711-s23.github.io/class_activities/ca_lecture_28.html