Mathematical Biostatistics Boot Camp 2: Lecture

Brian Caffo

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
Johns Hopkins University

September 30, 2013

Table of contents

Multiplicity

- After rejecting a χ^2 omnibus test you do all pairwise comparisons
- You conducted a study with 20 outcomes and 30 different combinations of covariates. You consider significance at all combinations.
- You compare diseased tissue versus normal tissue expression levels for $20k$ genes
- You compare rest versus active at $300k$ voxels in an fMRI study

Multiplicity

- Performing two α-level tests:
  H_0^1 versus H_a^1 and H_0^2 versus H_a^2
  E_1 Reject H_0^1 and E_2 Reject H_0^2

$$
\begin{aligned}
FWE & P(\text{one or more false rej} \mid H_0^1, H_0^2) \\
& = P(E_1 \cup E_2 \mid H_0^1, H_0^2) \\
& = P(E_1 \mid H_0^1, H_0^2) + P(E_2 \mid H_0^1, H_0^2) \\
& - P(E_1 \cap E_2 \mid H_0^1, H_0^2) \\
& \leq P(E_1 \mid H_0^1, H_0^2) + P(E_2 \mid H_0^1, H_0^2) \\
& = 2 \times \alpha
\end{aligned}
$$

Result : The **familywise error rate** for k hypotheses tested at level α is bounded by $k\alpha$

Proof

E_i - false rejection for test i
All probabilities are conditional on all of the nulls being true

$$\begin{aligned}
FWE &= P(\text{one or more false rej}) \\
&= P(\cup_{i=1}^k E_i) \\
&= P\left\{E_1 \cup (\cup_{i=2}^k E_i)\right\} \\
&\leq P(E_1) + P(\cup_{i=2}^k E_i) \\
&\vdots \\
&\leq P(E_1) + P(E_2) + \ldots + P(E_k) \\
&= k\alpha
\end{aligned}$$

Other direction

- The *FWE* is no larger than $k\alpha$ where k is the number of tests
- The *FWE* is no smaller than α

$$P(\cup_{i=1}^{k} E_i) \geq P(E_1) = \alpha$$

- The lower bound is obtained when the E_i are identical $E_1 = E_2 = \ldots = E_k$
- **Bonferoni's** tests each individual hypothesis at level $\alpha^* = \alpha/k$
  - The *FWE* is no larger than $k\alpha^* = k\alpha/k = \alpha$
  - The *FWE* is no smaller than α/k

Bonferoni's procedure

If α^* is small and the tests are independent, then the upper bound on the *FWE* is nearly obtained

$$
\begin{aligned}
FWE &= P(\text{one or more false rej}) \\
&= 1 - P(\text{no false rej}) \\
&= 1 - P(\cap_{i=1}^{k} \bar{E}_i) \\
&= 1 - (1 - \alpha^*)^k \\
&\approx 1 - (1 - k\alpha^*) \\
&= k\alpha^* = \alpha
\end{aligned}
$$

Scratch work

Recall the approximation for α^* near 0

$$\frac{f(\alpha^*) - f(0)}{\alpha^* - 0} \approx f'(0)$$

hence

$$f(\alpha^*) \approx f(0) + \alpha^* f'(0)$$

In our case $f(\alpha^*) = (1 - \alpha^*)^k$ so $f(0) = 1$

$f'(\alpha^*) = -k(1 - \alpha^*)^{k-1}$ so $f'(0) = -k$

Therefore $(1 - \alpha^*)^k \approx 1 - k\alpha^*$

Notes

- For Bonferoni's procedure $\alpha^* = \alpha/k$ so will be close to 0 for a large number of tests
- When there are lots of tests that are (close to) independent, the upper bound on the *FWE* used is appropriate
- When the test are closely related, then the *FWE* will be closer to the lower bound, and Bonferoni's procedure is conservative
- Is the familywise error rate always the most appropriate quantity to control for?

FDR

- The **false discovery rate** is the proportion of tests that are falsely declared significant
- Controlling the FDR is less conservative than controlling the FWE rate
- Introduced by Benjamini and Hochberg

Benjamini and Hochberg procedure

1. Order your k p-values, say $p_1 < p_2 < \ldots < p_k$
2. Define $q_i = kp_i/i$
3. Define $F_i = min(q_i, \ldots, q_k)$
4. Reject for all i so that F_i is less than the desired FDR

Note that the F_i are increasing, so you only need to find the largest one so that $F_i < FDR$

Example

1st 10 of 50 SNPs (Rosner page 581)

Gene	i	p_i	$q_i = kp_i/i$	F_i
30	1	$<.0001$	.0035	.0035
20	2	.011	.28	.16
48	3	.017	.28	.16
50	4	.017	.22	.16
4	5	.018	.18	.16
40	6	.019	.16	.16
7	7	.026	.18	.18
14	8	.034	.21	.21
26	9	.042	.23	.23
47	10	.048	.24	.24

Example

- Bonferoni cutoff $.05/50 = .001$; only the first Gene is significant
- For a FDR of $0 - 15\%$; only the first Gene would be declared significant
- For a FDR of $16 - 20\%$, the first 7 would be significant