Date: Tuesday, April 2, 2019

Example of jelly beans causing acne:

- 20 colors investigated
- $\alpha = 0.05$

H_{0j} : Color j jelly bean does not lead to acne
H_{Aj} : It does

$$\mathbb{P}(\text{reject at least one null hypothesis} \mid \text{all were true})$$

$$= \mathbb{P}(FD \geq 1)$$

↳ false discovery

FWER

$$= \mathbb{P}\left(\bigcup_{i=1}^{20} \{ \text{reject } H_{0j} \mid H_{0j} \text{ true} \} \right)$$

(Boole's ineq.)

$$\leq \sum_{j=1}^{20} \mathbb{P}(\text{reject } H_{0j} \mid H_{0j} \text{ true})$$

$$= \sum_{j=1}^{20} 0.05 = 20(0.05) = 1.$$

↳ Without controlling for FWER, we can expect (with high probability) to reject at least one H_{0j} ↳ incorrectly

⇒ we'll find that some color jelly bean causes acne.

$$\mathbb{P}(FD \geq 1) = 1 - \mathbb{P}(FD = 0) = 1 - (1 - 0.05)^{20}$$

(Because of indep.)

$$1 - (0.95)^{20}$$

$$0.64$$

* So we should correct for multiple comparisons!

First strategy: Control for FWER.

Method: Bonferroni's correction.

- Reject H_{0j} when the p-value of the test $p_j \leq \frac{\alpha}{m}$

This will guarantee that FWER $\leq \alpha$!

Jelly bean example:

$$FWER = 1 - \left(1 - \frac{0.05}{20}\right)^{20}$$

$$= 0.0488 \checkmark$$

But now our p-value cut-off is drastically reduced

⇒ leads to rarely rejecting

⇒ leads to increased type II error for each test

⇒ leads to low power for each test.

Typically, we refer to Bonferroni as a conservative strategy because we will have few rejections.

* There are applications where being conservative on rejection is desired.

Ex: Malware detection
Pharmaceutical trials
Gene activation

Question: How can we specify a multiple comparisons strategy that is less conservative? (and more powerful?)

1) We can try alternative methods to control FWER, but in general these are still conservative.

2) We can look at alternative criterions to FWER. A popular one is the false discovery rate, FDR.

Recall: FWER — $\mathbb{P}$(having at least one FD in m tests)

↳ To slacken the restriction of FWER, we can allow for a few more FD's.

⇒ FDR does this!

Recall: R = # of rejection from m tests
V = # of false rejection.

$\frac{V}{R}$ = The rate of false discovery but note that V is an unknown random quantity.

$$\Rightarrow \text{FDR} = \mathbb{E}\left[\frac{V}{R} \,\middle|\, R \geq 1\right] = \frac{\mathbb{E}[V \mid R \geq 1]}{\max\{R, 1\}}$$

We'd like to control FDR!

i.e. to ensure that FDR $\leq \alpha$. (1)

↳ Thankfully, in 1995 Benjamin and Hochberg developed a strategy that guarantees (1)

↳ The paper "Controlling the false discovery rate:..." is the most cited statistical paper out there (53,710 as April 2nd, 2019) compared to the Lasso paper (from 1996 w/ 27549).

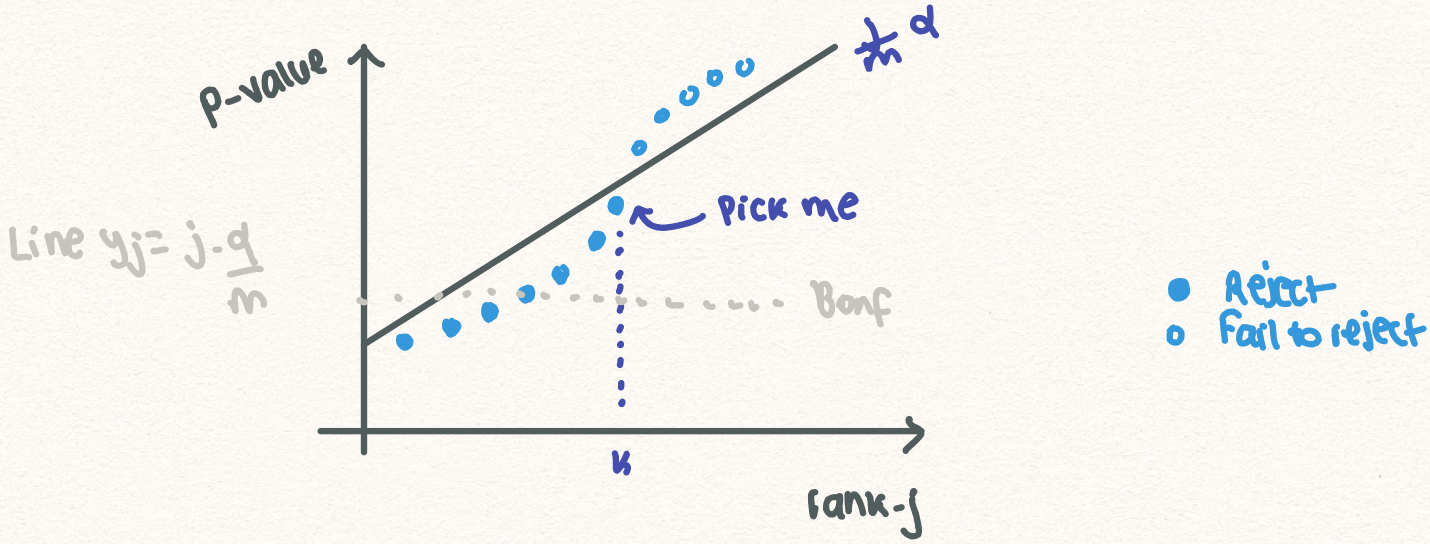
B-H step-up Procedure

Setup: 1) Orders p-values from smallest to largest

$P_{(1)} \le P_{(2)} \le \dots \le P_{(m)}$ where $P_{(j)}$ = j^{th} smallest p-value

2) Find the largest k such that

$P_{(k)} \le \frac{k}{m}\alpha$ and $P_{(j)} \le \frac{j}{m}\alpha$ for all $j = 1, \dots, k$



3) Reject all hypothesis $j = 1, \dots, k$

Notes:

- As the rank of the p-value increases, the cutoff increases linearly.

Ex: · $P_{(1)}$ is rejected if $\le \frac{\alpha}{m}$ (Bonferroni!)

· $P_{(m)}$ is rejected if $\le \alpha$ (no adjustment!)

Key Property: If the hypothesis tests are independent, then running the set-up procedure guarantees FDR $\le \alpha$.

If not independent, this holds approx. as $m \to \infty$

In class - excersice:

- Simulate 1,000 U(0,1) → treat as p-values

  1) Reject w/ no corr
  2) Bonferroni
  3) B-H
  * rejections

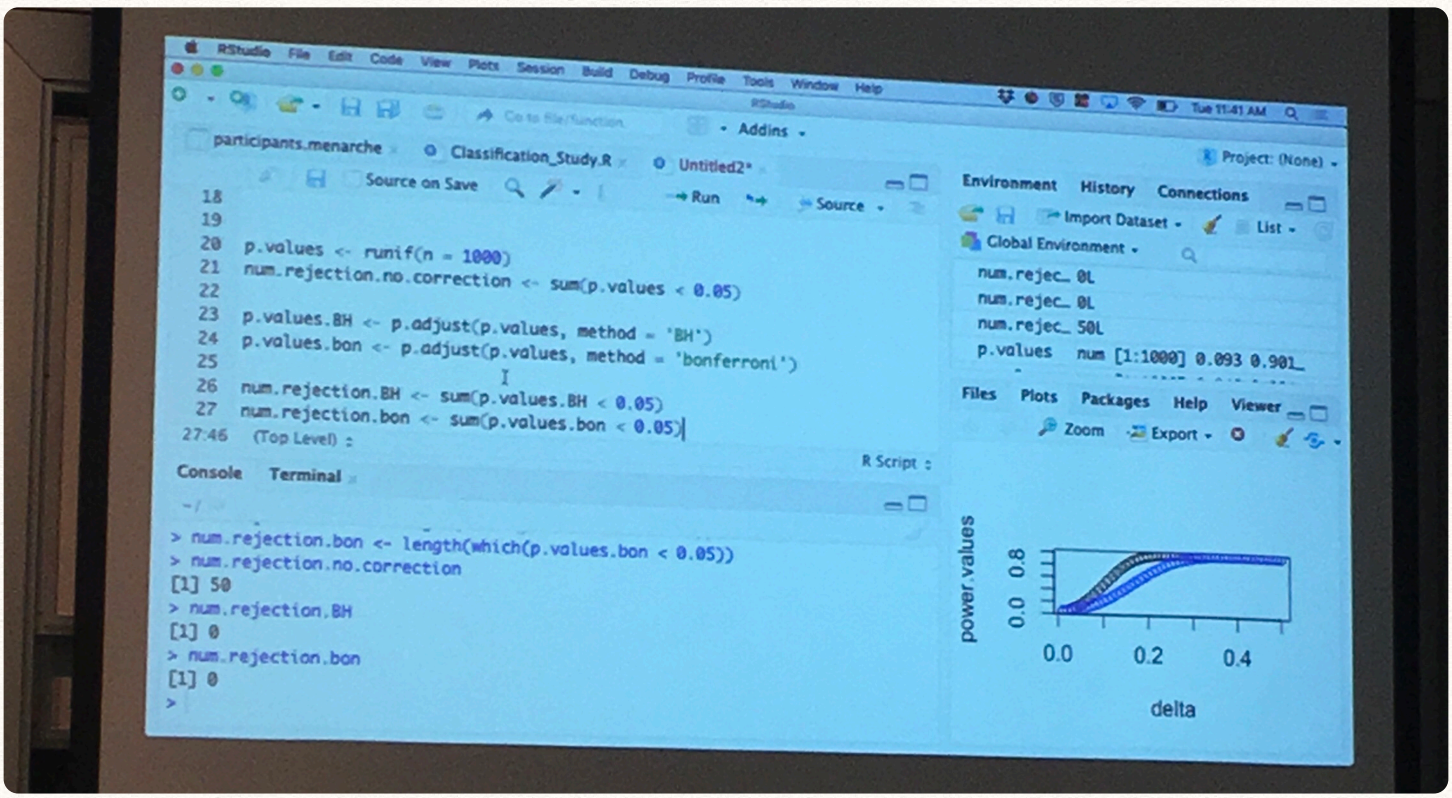
Note: Can approximate FDR $= \frac{\mathbb{E}[V \mid R \geq 1]}{\max\{R, 1\}}$ using Monte Carlo simulation :)

* Once we run step-up (or Bonferroni) we can calculate the observed false discovery proportion (fdp) as fdp $= \frac{V}{R}$ ← observed if we know which hypoth are true.

* If H_0 is true, one can show that p-val $\sim U(0,1)$

* Bonferroni, a corrected p-value will be $m \cdot$ p-value (if $\leq \alpha$, reject)

* In step up, it will be for the ith smallest p-value
$P_{(i)}^{(new)} = \frac{m\, P_{(i)}}{i}$ (if $\leq \alpha$, reject)



* How do FDR & FWER compare?

1) If all H_{0j} are true, FDR = FWER

2) In general, FDR $\leq$ FWER $\leq \alpha$

$\Rightarrow$ If we control FWER, we also control FDR.

FDR $\approx \mathbb{P}(V > 1) \leq \mathbb{P}(V \geq 1)$