Feynman Liang
STAT 210A
Lecture #28
12/1/2016

1 Simultaneous vs Selective Inference

Last time: Solve all inference problems at once, guarantee 0 type I errors (w.p. $1-\alpha$)
Logic: Then, ok to cherrypick "interesting" results

Example 1.1. $X_i = \mu_i + \varepsilon_i$, $\varepsilon_i \overset{\text{iid}}{\sim} N(0,1)$, $i = 1, \dots, n$

1. Select $i^* = \arg\max X_i$
2. Report a CI for μ_{i^*}. Note μ_{i^*} not necessarily equal to $\max_i \mu_i$.

1.1 Simultaneous approach

Construct intervals $C_i = X_i \pm z_{\tilde{\alpha}_n/2}$ where

$$\tilde{\alpha}_n = 1 - (1-\alpha)^{1/n} \approx \alpha/n \tag{1.1}$$
$$z_{\tilde{\alpha}_n/2} \approx \sqrt{2\log n} \tag{1.2}$$

Report $C_{i^*(X)(X)}$.

$$\mathbb{P}_\mu[\mu_{i^*} \notin C_{i^*}] \leq \mathbb{P}_\mu[\mu_i \notin C_i, \text{ any } i] = \alpha \tag{1.3}$$

Two issues:

1) We expect $X_{i^*} > \mu_{i^*}$ So why inflate symmetrically?
2) What if $\mu_1 \gg \max_{i>1} \mu_i$
   $\mathbb{P}[i^* = 1] \approx 1$
   $\mathbb{P}[\mu_{i^*} \notin C_{i^*}] \approx P[\mu_1 \notin C_1] \approx \alpha/n$

1.2 Selective approach

Idea: Design C_i to guarantee

$$\mathbb{P}_\mu[\mu_i \notin C_i \mid i^* = i] = \alpha \tag{1.4}$$

Then we have

$$\mathbb{P}_\mu[\mu_{i^*} \notin C_{i^*}] = \mathbb{E}_\mu[\mathbb{P}_\mu[\mu_i^* \notin C_{i^*} \mid i^*]] = \alpha \tag{1.5}$$

The conditional distribution

$$P(X \mid i^*(x) = i) \propto e^{-\frac{1}{2}\|x-\mu\|^2} 1_{\underbrace{x_i > \max_{j \neq i} x_j}_{A_i}} \tag{1.6}$$

$$\propto e^{\mu' x} e^{-\|x\|^2/2} 1_{x \in A_i} \tag{1.7}$$

$$= e^{\mu_i x_i + \mu'_{-i} x_{-i}} e^{-\|x\|^2/2} 1_{x \in A_i} \tag{1.8}$$

UMPU test of $H_0 : \mu_i = c$ condition on x_i

Reduced the problem to same case as when we had nuisance parameters.

TODO: Fig 28.1

Note: Interval for μ_i only defined as A_i.

Suppose $\mu_1 \gg \max_{i>1} \mu_i$.

Then $X_1 \gg \max_{i>1} X_i$ (whp) and conditional inference $\approx$ marginal inference. So $C_{i^*} \overset{whp}{=} C_1 \approx X_1 \pm z_{\alpha/2}$.

2 False Discovery Rate (FDR)

In ~~1995~~1988 Benjamini & Hochberg proposed controlling a more liberal error criterion for multiple testing: *False Discovery Rate*.

Idea: Test 500k hypothesis and make 100 rejections, of which 5 are false.

Let $V = \#$ false rejections and $R = \#$ total rejections.

Definition 2.1. The *False Discovery Proportion (FDP)*

$$FDP = \frac{V}{R} \quad \left[\frac{0}{0} = 0\right] \tag{2.1}$$

The FDP is a random variable, so the *False Discovery Rate (FDR)*

$$FDR = \mathbb{E}[FDP] \tag{2.2}$$

Proposal: control $FDR \leq \alpha$

2.1 BH Procedure

Setup: $p_1, \ldots, p_n$ p-values for $H_{0,1}, \ldots, H_{0,n}$ independent [uniform if null].

Let $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}$, $H_{0,(i)}$ = hypothesis corresponding to $p_{(i)}$.

Select $R = \{r : p_{(r)} \leq \frac{\alpha r}{n} \text{ (or 0)}\}$.

Reject $H_{0,(1)}, \ldots, H_{0,(R)}$ (or none if $R = 0$).

TODO: Fig 28.2

Intuition: Consider the rule rejecting all $p_i \leq t$ for fixd $t \in [0,1]$.

$$FDP(t) = \frac{\#\{i \in \mathcal{H}_0 : p_i \leq t\}}{1 \vee \#\{i : p_i \leq t\}} = \frac{V_t}{1 \vee R_t} \tag{2.3}$$

But $\mathbb{E}V_t = t|\mathcal{H}_0 rvert \leq tn$ and

$$\widehat{FDP}(t) = \frac{nt}{1 \vee R_t} \tag{2.4}$$

Idea: Could take $t^* = \sup\{t : \widehat{FDP}(t) \leq \alpha\}$. Reject all $H_{0,i}$ with $p_i \leq t^*$.

Remark 2.2. If R is BH procedure, then $p_{(R)} \leq t^* \leq p_{(R+1)} \implies$ BH-procedure equivalent to "t^* procedure"

Proof. Consider $t = p_{(i)}$.

$$\widehat{FDP}(p_{(i)}) \leq \alpha \iff \frac{np_{(i)}}{i} \leq \alpha \iff p_{(i)} \leq \alpha \tag{2.5}$$

Also $\widehat{FDP}(t) \uparrow$ in t between order statistics (jumps at order statistics)

TODO: Fig 28.3 $\square$

Theorem 2.3 (Storey, Taylor, & Siegmund). *BH-procedure controls FDR.*

Define $M_t = \frac{V_t}{nt}$. M_t is a MG as t goes from $1 \to 0$ wrt $\mathcal{F} = \sigma(p_i : i \notin \mathcal{H}_0$ or $p_i > t)$

For $s < t$, $V_s \mid \mathcal{F}_t \sim Binom(V_t, \frac{s}{t})$.

$$\mathbb{E}[M_s \mid \mathcal{F}_t] = \frac{1}{ns}\mathbb{E}[V_s \mid \mathcal{F}_t] \tag{2.6}$$

$$= \frac{1}{ns}V_t \cdot \frac{s}{t} \tag{2.7}$$

$$= \frac{1}{\cancel{ns}}(M_t \cancel{nt})\frac{\cancel{s}}{\cancel{t}} = M_t \tag{2.8}$$

t^* is a stopping time wrt $(\mathcal{F}_t)$.

$$FDR = \mathbb{E}[FDP(t*)] = \mathbb{E}[\widehat{FDP}(t^* M_{t^*})] = \mathbb{E}[\alpha M_{t^*}] = \alpha \mathbb{E}[M_1] = \alpha \frac{|\mathcal{H}_0|}{TODO : ??} \leq \alpha \tag{2.9}$$