Lecture 6. Multiple comparison problem.

store $\boxed{1, 2, \ldots\ldots M}$

H_0^i : the measure is not eff in store i
H_1^i : — // — is effect. in store i

$\boxed{\text{Bad practice}}$

α "=" $P(H_0 \text{ rej} \mid H_0)$" $= 0.05$

test all H_0^i with this $\alpha = 0.05$

H_0^1 not rej
H_0^2 not rej
H_0^3 rej → that works!
H_0^4 not rej
H_0 rej → that works!
H_0^M not rej

If you say that actually $\alpha = 0.05$ then you are wrong!

Ex

$100: H_0^1, \ldots\ldots H_0^{100}$
↳ Data №1 ↳ Data №100 $\boxed{\text{indep.}}$

I am using $\alpha = 0.05$ for each H_0^i
Imagine that All H_0^i are true!

a) $P(\text{I will reject at least one } H_0^i) \stackrel{?}{=} 1 - (1-\alpha)^{100} \approx$
≈ 0.994

b) $E(R) ? \stackrel{?}{=}$
R – number of rejected H_0^i

$R \sim Bin(n = 100, p = \alpha)$ $E(R)$ "=" np "=" $100 \cdot 0.05 = 5$

One more example (bad practice)

$$y_i = \beta_1 \cdot 1 + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i3} + \ldots \beta_{10000} \cdot x_{i10000} + u_i$$

$\to$ look at the signif of $\hat{\beta}_j$ on $\alpha = 0.05$

The remedies !!

Simplest one: Bonferonni correction !!

the number of tested H_0^i

$\downarrow$

$M = U + T + V + S$

conc. / TRUTH	do not rej H_0	rej H_0
H_0	!! U	I type V
H_0^C	II type T	!! S'

$V + S = R$ - the number of rejected H_0^i

V - the number of wrongly rejected H_0^i

def

FWER - family wise error rate $= P(V > 0)$

Goal: A procedure that guar. FWER $\leq \alpha$

Simplest sol-n. Bonferonni correction.

Test each individual H_0^i with $\alpha_i = \frac{\alpha}{M}$.

Proof

$P(A_1 \cup A_2 \cup A_3) \leq P(A_1) + P(A_2) + P(A_3)$

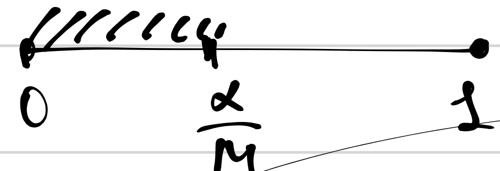
p_i - p-value in test of H_0^i

$$\underline{FWER} = P(V>0) =$$
$$= P\Big(\bigcup \{p_i < \tfrac{\alpha}{M}\}\Big) \leq$$
for all H_0^i that are true

$$\leq \sum_{H_0^i \text{ is true}} P\Big(p_i < \tfrac{\alpha}{M}\Big) = \sum_{H_0^i \text{ is true}} \frac{\alpha}{M} \leq M \cdot \frac{\alpha}{M} = \alpha$$

$\boxed{\text{OLD FACT}}$ $\boxed{\text{Under } H_0^i \quad \text{p-value} \quad p_i \sim U[0;1]}$

$\boxed{!}$



$p_i \sim U[0;1]$

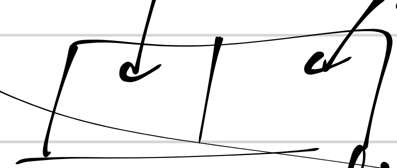
des. is ok.

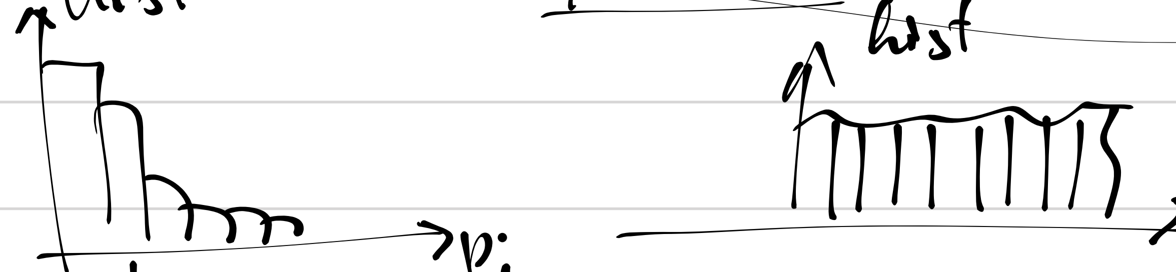
A/A - testing.

A - control [old]

A - experim. [old]

$\times 100$





hist

hist

$\to p_i$ $\to$ I may start real A/B test

design is bad !!

Proof. Under H_0 p-value $\sim U[0;1]$ distr.

Proof for specific ex. (the statement is true in general)

$X_1 \dots X_n \sim$ iid $N(\mu; 1)$

that's never known !!

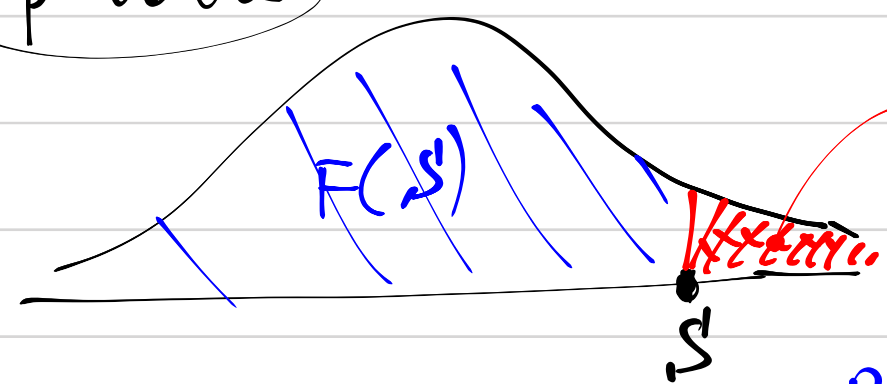
$H_0: \mu = 5$

$H_A: \mu > 5$

$$S = \frac{\bar{X} - 5}{\sqrt{\frac{1}{n}}} \overset{H_0}{\sim} N(0;1)$$

$\frac{\sigma^2}{n}$

p-value



$F(S)$

p-value

$$P(\text{p-value} < 0.05) = P(S > 1.65) =$$
$$= 0.05 \;!!$$

$$P(\text{p-value} < t) =$$
$$= P(1 - F(S) < t) =$$
$$= P(F(S) > 1-t) = P(S > F^{-1}(1-t)) =$$
$$= 1 - P(S < F^{-1}(1-t)) = 1 - F(F^{-1}(1-t)) = t \;!!$$

Holm-Bonferroni correction. [use this instead of Bonf corr]

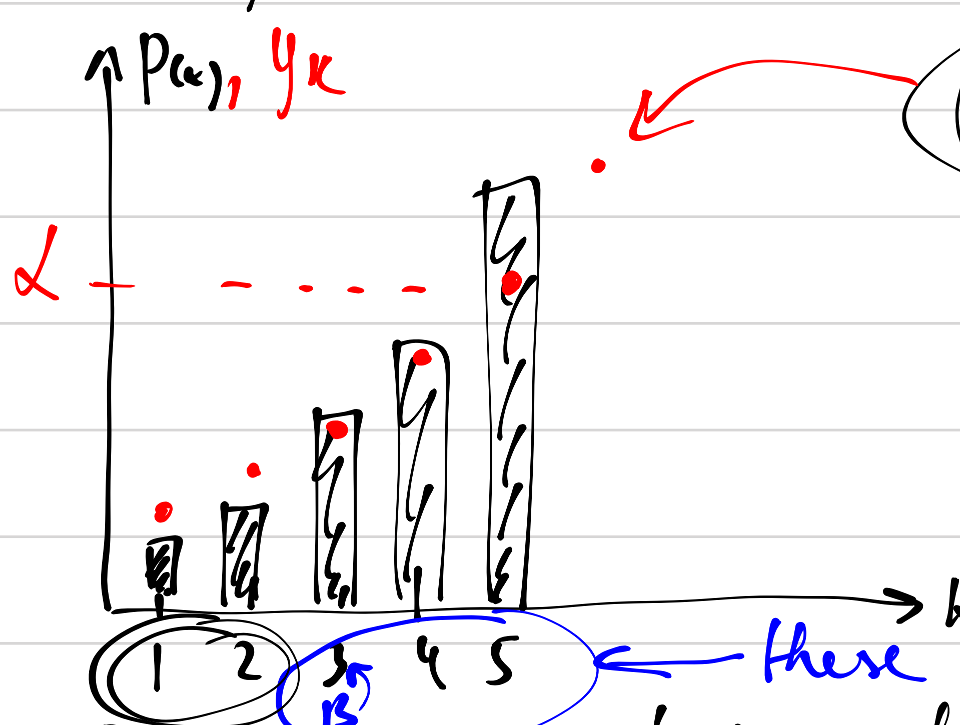
Aim: FWER $\le \alpha$

Alg. 1) Sort p-values: $p_{(1)} \le p_{(2)} \le \dots \le p_{(M)}$

2) Draw them with the line $y_k = \dfrac{\alpha}{M+1-k}$

$k \uparrow$, $M+1-k \downarrow$, $y(k) \uparrow$

M - the number of tested H_0^i



these hypo are not rejected

are rejected

3) Find first k such that $p_{(k)} \ge y_k$

$B = \min\{k \mid p_{(k)} \ge y_k\}$

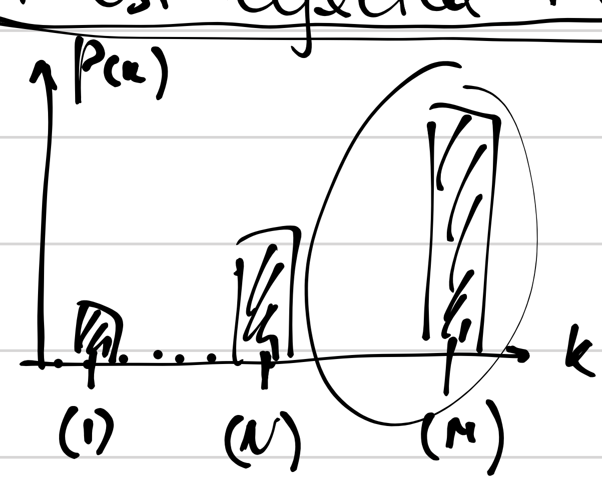
reject hypo (1), (2) .... (B-1)

do not reject (B), (B+1), (B+2) .... (M)

M - is the total number of hypo tested

M_0 - is the number of true H_0^i $M_0 \le M$

Step! First rejected true hyp has $p_{(N)} < \dfrac{\alpha}{M_0}$



To the left of (N) all hypo H_0^i are false

If (N) is rej $\Rightarrow$ (N-1), (N-2) ... (1) are also reject.

(M-N) hypo to the right of hyp (N)

$\underbrace{(M-N) + 1}_{} \ge M_0$; 1 (N)

$\boxed{p_{(N)}} < \dfrac{\alpha}{M+1-N} \le \dfrac{\alpha}{M_0}$

$$p_{(i)} \le \frac{\alpha}{M_0}$$

Step 2.

$$FWER = P(V>0) \le P\left(\text{At least one } p_{(i)} \le \frac{\alpha}{M_0}\right) \le$$

$$= P\left(\bigcup_{H_0^i \text{ is true}} p_{(i)} \le \frac{\alpha}{M_0}\right) \le$$

$$\le \sum_{H_0^i \text{ is true}} P\left(p_{(i)} \le \frac{\alpha}{M_0}\right) = M_0 \cdot \frac{\alpha}{M_0} = \alpha.$$

$p_{(i)} \sim U[0;1]$