!! L8.

$\boxed{\text{Don't PANIC !!}}$

Multiple comparisons.

	H_0 is not rej	H_0 is rej
H_0 TRUE	U	V
H_0 FALSE	T	S

$M = U + V + T + S$

$M_0 = U + V$

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

$\boxed{\text{goal : FWER} \le \alpha}$

Second goal:

FDR = false discovery rate.

$FDR = E(V/(V+S))$ (not ok)

$FDR = E\left(\frac{V}{\max(V+S, 1)}\right)$ (correct)

theorem: FWER $\geq$ FDR

$I = \begin{cases} 1 & \text{if } V \geq 1 \\ 0 & \text{if } V = 0 \end{cases}$

proof $FWER = E(I(V \geq 1))$

$E(I(V \geq 1))$ vs $E\left(\frac{V}{\max(V+S, 1)}\right)$

$I(V \geq 1)$ vs $\frac{V}{\max(V+S, 1)}$

case A: $V=0$ $LHS = 0 \geq RHS = 0$

case B: $V \geq 1$ $LHS = 1 \geq RHS \leq 1$

FWER $\geq$ FDR

Benjamin-Hochberg procedure.

control FDR

goal: FDR $\le \alpha$

Step 1

Test: $H_0^1, H_0^2, \ldots\ldots H_0^M$

we obtain p-values

we sort them:

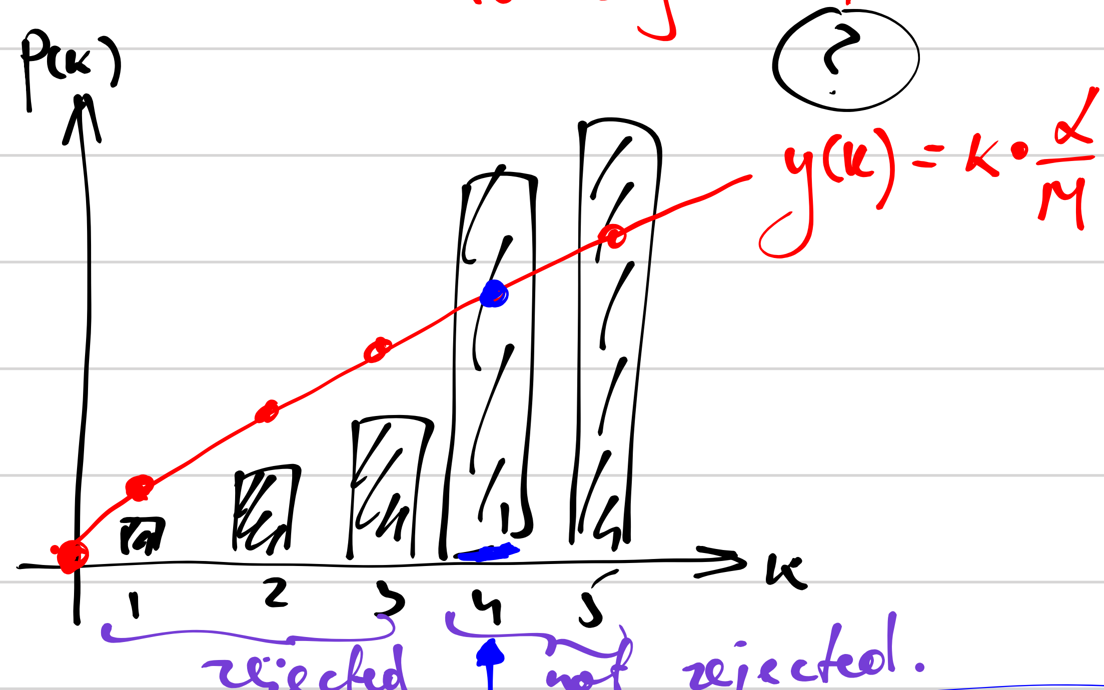
$$p_{(1)} \le p_{(2)} \le \ldots\ldots \le p_{(M)}.$$

idea. too small | too big

$\underbrace{\qquad\qquad}_{H_0 \text{ rej}}$ | $\underbrace{\qquad\qquad}_{H_0 \text{ not reject}}$

?

Step 2.



F = Find the first $p_{(k)}$ above the line $y(k)$.

Step 3

Do not reject $H_0^F, H_0^{F+1}, \ldots\ldots H_0^M$

Reject $H_0^1, H_0^2, \ldots H_0^{F-1}$

---

proof is more involved !!

$$FDR = E\left(\frac{V}{\max(1, V+S)}\right)$$

		H_0 rej
H_0 TRUE		V
		S

general idea.

$$\frac{V}{\max(1,V+S)} = \sum_{i \in I_{H_0}} \frac{V_i}{\max(1,V+S)} = (*)$$

H_0^1	H_0^2	. . .	–	H_0^M
F	T		T	F

M null-hypo in total.

I_{H_0} = $\boxed{\text{the set of indices of true } H_0}$

$I_{H_0} = \{2, 4, \dots 9\}$

$V_i = \begin{cases} 1 & \text{if the hypo } i \text{ was rejected.} \\ 0 & \end{cases}$

$$* = \sum_{i \in I_{H_0}} X_i \qquad X_i = \frac{V_i}{\max(1,V+S)}$$

$\boxed{\text{Claim}}$: for $i \in I_{H_0}$ $\left(E(X_i) = \frac{\alpha}{M} \right)$ (ind. pval)

$$FDR = E\left(\frac{V}{\max(1,V+S)}\right) = E\left(\sum_{i \in I_{H_0}} X_i\right) =$$

$$= \sum_{i \in I_{H_0}} E(X_i) = M_0 \cdot \frac{\alpha}{M} = \frac{M_0}{M} \cdot \alpha \le \alpha.$$

!!

Ex.

three hypo H_0^A, H_0^B, H_0^C — indep. data

$[\text{All } H_0^A, H_0^B, H_0^C \text{ are true}]$

$p_{(1)}, p_{(2)}, p_{(3)}$

a) pdf for each sorted $p_{(k)}$?

b) $P\left(p_{(1)} \geq \frac{\alpha \cdot 1}{3}\right) \stackrel{?}{=}$ $\qquad y(k) = \frac{\alpha \cdot k}{M}$, $M = 3$

c) $P\left(p_{(1)} < \frac{\alpha \cdot 1}{3}, p_{(2)} \geq \frac{\alpha \cdot 2}{3}\right) = ?$ $P(\text{ex. one null will be rej})$

d) $P\left(p_{(1)} < \frac{\alpha}{3}, p_{(2)} < \frac{\alpha \cdot 2}{3}, p_3 < \frac{\alpha \cdot 3}{3}\right)$?

e) FDR ?

a) $p_{(A)}, p_{(B)}, p_{(C)} \sim U[0;1]$

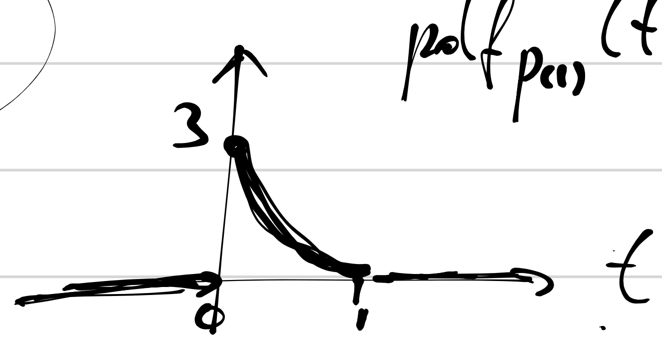
$p_{(1)} = \min(p_{(A)}, p_{(B)}, p_{(C)}) \not\sim U[0;1]$

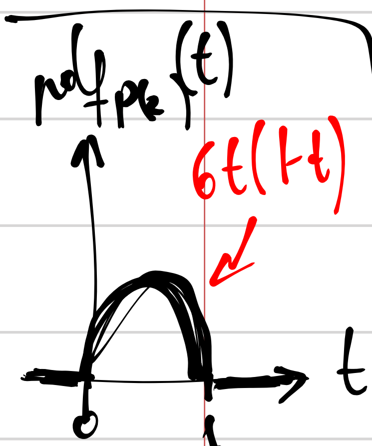
$p_{(1)}$ cdf: $P(p_{(1)} \leq t) = 1 - P(p_{(1)} > t) =$

$= 1 - P(p_{(A)} > t, p_{(B)} > t, p_{(C)} > t) =$

$= 1 - (1-t)^3$

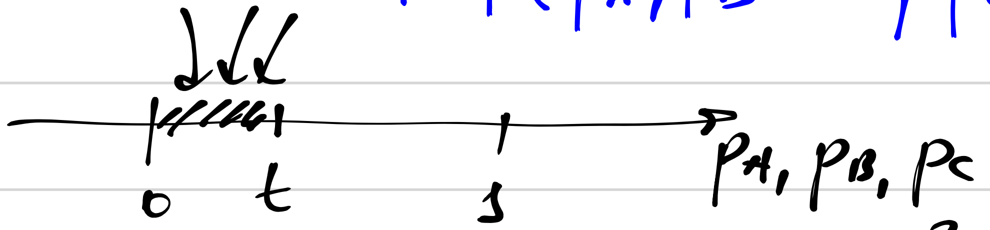
pdf $= 3 \cdot (1-t)^2$

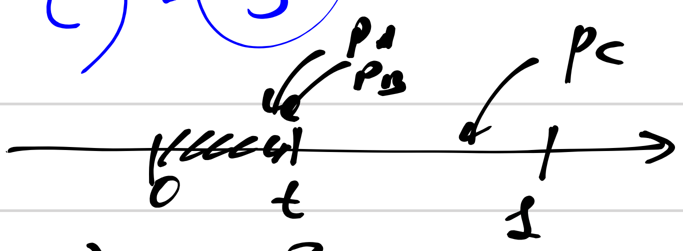




$p_{(2)}$ cdf $P(p_{(2)} \leq t) = P(p_A, p_B, p_C \leq t) +$

$+ P(p_A, p_B \leq t, p_C > t) \cdot 3 =$



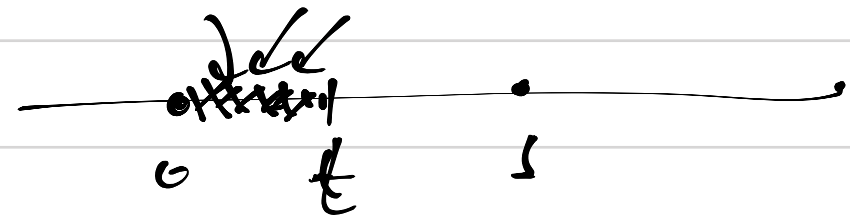


$= t^3 + 3 \cdot t^2 \cdot (1-t) = 3t^2 - 2t^3$

pdf $= 6t - 6t^2 = 6t(1-t)$

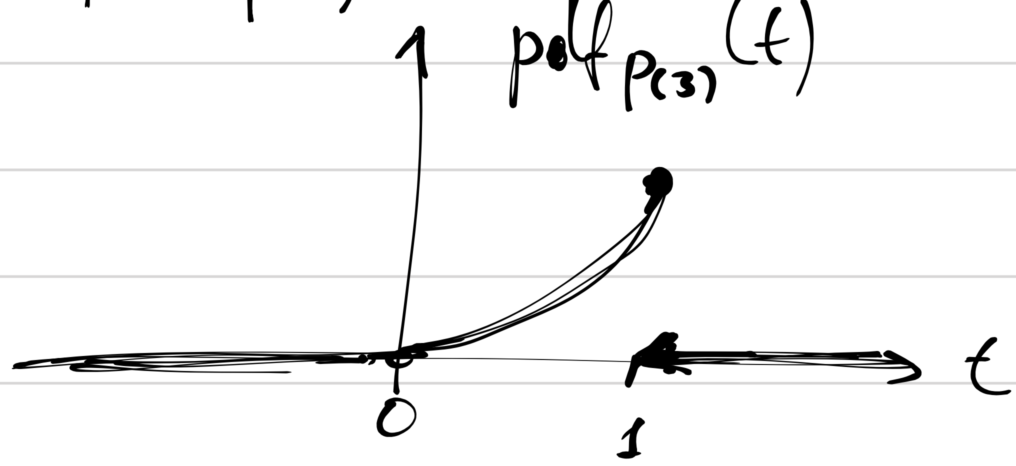
$p_{(3)} = \max(p_A, p_B, p_C)$

cdf. $P(p_{(3)} \le t) = P(p_{(A)} \le t, p_{(B)} \le t, p_{(C)} \le t) =$

$=$  $=$

$= t^3$

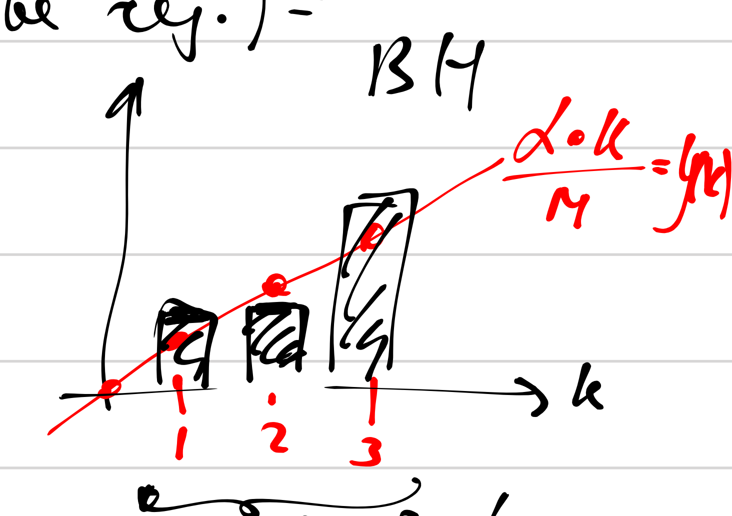
pdf $p_{(3)} = 3t^2$ on $[0;1]$



$\text{cdf}_{p_{(1)}}(t) = 1-(1-t)^3$

b) $P(\text{all three will not be rej.}) =$

$P\left(p_{(1)} \ge \frac{\alpha\cdot 1}{3}\right) =$

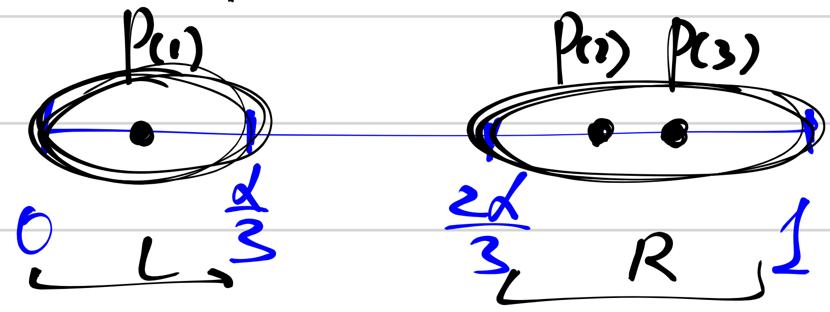


$= 1 - \text{cdf}_{p_{(1)}}\left(\frac{\alpha}{3}\right) =$

$= 1 - \left(1-\left(1-\frac{\alpha}{3}\right)^3\right) = \left(1-\frac{\alpha}{3}\right)^3$

p_A $\quad p_B$, p_C ~ 1

c)



$P(\text{ex. one null hypo will be rej}) =$

$= P\left(p_{(1)} < \frac{\alpha}{3},\ p_{(2)} \ge \frac{2\alpha}{3}\right) = \left(\frac{\alpha}{3}\right)\cdot\left(1-\frac{2\alpha}{3}\right)^2 \times 3$

$P(p_A \in L, p_B \in R, p_C \in R) = \frac{\alpha}{3}\cdot\left(1-\frac{2\alpha}{3}\right)^2\cdot 3$

$P(p_A \in R, p_B \in L, p_C \in R)$

$P(\text{— // —},\ p_C \in L)$

b) P(no null hypo is reject) $= (1-\frac{\alpha}{3})^3$

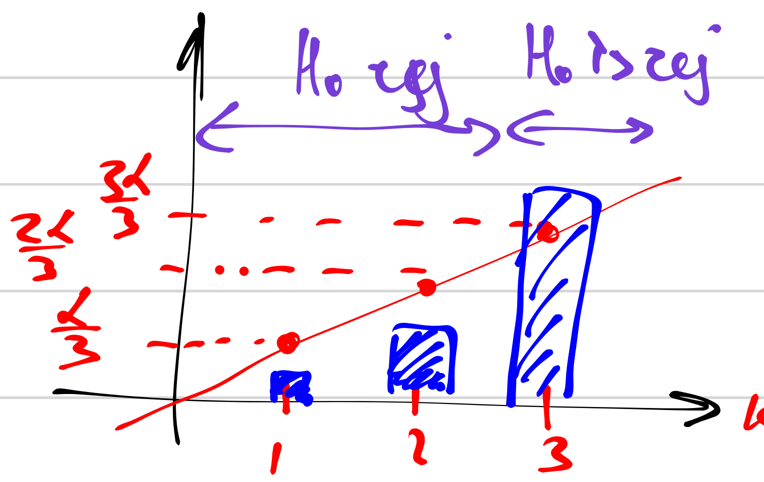
$P(V=0) = (1-\frac{\alpha}{3})^3$

c) $P(V=1) = \alpha(1-\frac{2\alpha}{3})^2$

d) $P(V=2) =$

$P(V=3) =$

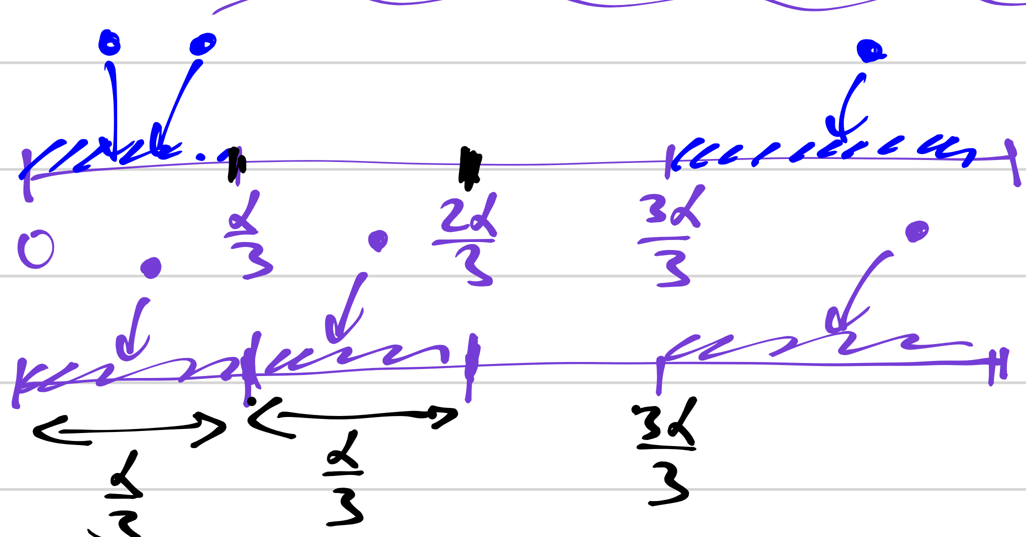
$P(V=2) = ?$



$$P(V=2) = P\left(p_{(1)} < \frac{\alpha}{3},\ p_{(2)} < \frac{2\alpha}{3},\ p_{(3)} \geq \alpha\right) =$$

type A:

type B



type A: $\left(\frac{\alpha}{3}\right)^2 \cdot \left(1-\frac{3\alpha}{3}\right)$

type B: $\frac{\alpha}{3} \cdot \frac{\alpha}{3} \cdot \left(1-\frac{3\alpha}{3}\right)$

$$= \left(\frac{\alpha}{3}\right)^2\left(1-\frac{3\alpha}{3}\right) \cdot (3+6)$$

e)

$$FDR = P(V=0)\cdot 0 + P(V=1)\cdot 1 + P(V=2)\cdot 1 + P(V=3)\cdot 1$$

$$FDR = E\left(\frac{V}{\max(V+S',1)}\right) =$$

$S'=0.$

		H_0 rej
H_0 TR	U	V
H_0 FALSE	T=0	S'=0

False discov.

Family wise error rate

check $FDR \leq \alpha$.

$$FDR = 1-P(V=0) = 1-(1-\frac{\alpha}{3})^3 =$$

$$= 1-\left(1-3\cdot\frac{\alpha}{3}+9\cdot\frac{\alpha^2}{9}-\frac{\alpha^3}{27}\right) \vee \alpha$$

$$-\alpha^2+\frac{\alpha^3}{27} \vee 0$$

$$-27+\alpha \vee 0 \qquad \alpha \leq 27$$

$FDR \leq \alpha$

BH p.