Γ $(\Rightarrow)$ $H_{(j)}$ is rejected $\Rightarrow$ $\bigcap_{i=1}^{n} H_{(i)}, \dots, \bigcap_{i=j}^{n} H_{(i)}$ are rejected $\Rightarrow$ $\varphi_{I^+_{(1)}} = 1, \dots, \varphi_{I^+_{(j)}} = 1$

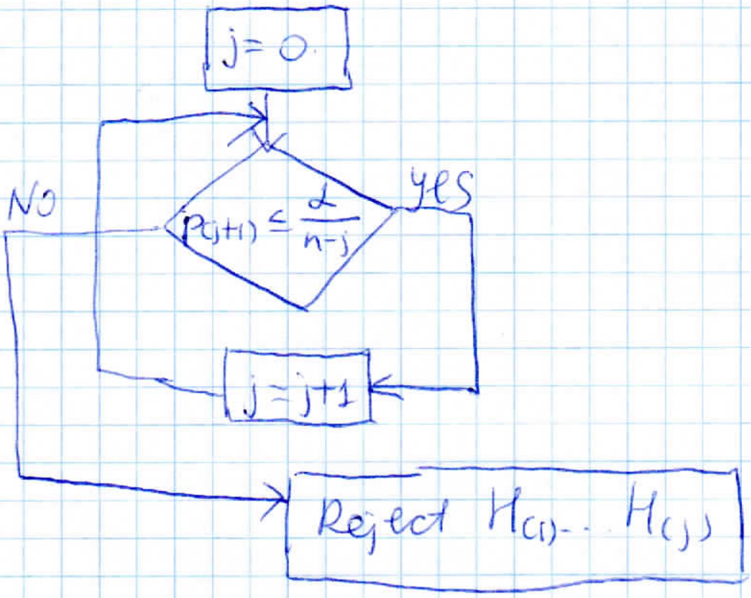
$(\Leftarrow)$ $\varphi_{I^+_{(1)}} = 1, \dots, \varphi_{I^+_{(j)}} = 1.$

$\varphi_{I^+_{(j)}} = 1 \Rightarrow \forall I : I \subset \{i_j, \dots, i_n\},\ i_j \in I : \varphi_I = 1$

We can take $I = \{i_j\} \Rightarrow \varphi_{i_j} = 1 \Rightarrow$

$\Rightarrow$ $H_{(j)}$ is rejected $\rfloor$

Procedure Closed Bonferroni:



Closing Simes   p-values are independent!

Simes test statistic:

$$\varphi_I = 1 \quad \text{iff} \quad \& \begin{cases} p_{(1,I)} \le \frac{\alpha}{|I|} \\ p_{(2,I)} \le \frac{2\alpha}{|I|} \\ \dots \\ p_{(|I|,I)} \le \frac{|I|\alpha}{|I|} = \alpha \end{cases}$$

Here we derive a simple procedure that is strictly more conservative than the closure of Simes. Our procedure will also control FWER under independence.

Lemma 2.

$$p_{(1)} \le p_{(2)} \le \dots \le p_{(j)} \le \dots \le p_{(j')} \le \dots \le p_{(n)}$$
$$H_{(1)} \quad H_{(2)} \quad\quad H_{(j)} \quad H_{(j')} \quad H_{(n)}$$
$$H_{i_1} \quad H_{i_2} \quad\quad H_{i_j} \quad H_{i_{j'}} \quad H_{i_n}$$

index set I suppose that:

(a) $i_j \in I$

(b) $\exists j' \ge j$ such that
$$p_{(j')} \le \frac{\alpha}{n-j'+1}$$

Then $\varphi_I = 1$ for the Simes test φ_I

⌈ Let $k := \max\{\ell : i_\ell \in I, \ \ell \le j'\}$. By (a) k exists and finite. Then

$$p_{(k)} \le p_{(j')} \le \frac{\alpha}{n-j'+1} \le \left| \begin{array}{l} |\{i_{j'+1} \dots i_n\}| = n-j' \\ \Rightarrow |\{i_{j'+1} \dots i_n\} \cap I| \le n-j' \end{array} \right| \le$$

$$\le \frac{\alpha}{1+|\{i_{j'+1} \dots i_n\} \cap I|} = |\text{by definition of } k| =$$

$$= \frac{\alpha}{|\{i_k, \dots, i_n\} \cap I|} \underset{*}{\le} \frac{|\{i_1 \dots i_k\} \cap I|}{|I|}\alpha$$

* ⌈ by definition $i_k \in I$.
Let $a = |\{i_1 \dots i_{k-1}\} \cap I|$, $b = |\{i_{k+1} \dots i_n\} \cap I|$
then $a+b+1 = |I| \Rightarrow |I| \le ab + a + b + 1 =$
$= (a+1)(b+1)$

$$\Rightarrow \frac{1}{b+1} \le \frac{a+1}{|I|} \ \lrcorner$$

By definition of Simes procedure $\varphi_I = 1$ ⌋

$\Rightarrow$ Reject $H_{(j)} \Leftrightarrow \exists j' \ge j$ such that $p_{(j')} \le \dfrac{\alpha}{n-j'+1}$

closure principle (closing Simes) $\Rightarrow$ this procedure controls FWER

This procedure is known as Hochberg's procedure.

Step - Down vs. Step - Up Procedures

<u>Holm</u>	<u>Hochberg</u>
$j = 0$	$j = n$
<u>while</u> $p_{(j+1)} \le \frac{\alpha}{n-j}$ <u>do</u>	<u>while</u> $p_{(j)} > \frac{\alpha}{n-j+1}$ <u>do</u>
$j = j+1$	$j = j-1$
<u>end</u>	<u>end</u>
Reject $H_{(1)} \dots H_{(j)}$	Reject $H_{(1)} \dots H_{(j)}$

<u>Example:</u>

$p_1 = \dots = p_n = \alpha \Rightarrow \forall j \quad p_{(j)} \le \frac{\alpha}{n-j} \Rightarrow$

$\Rightarrow$ Holm's procedure rejects nothing

$\Rightarrow j = n \quad p_{(n)} > \frac{\alpha}{n-n+1}$ - FALSE

$\Rightarrow$ Hochberg procedure rejects everything.

<u>In general</u>, step-up procedures can be more powerful than step-down procedures.

<u>Holm</u>

$j = 1$

<u>while</u> $p_{(j)} \le \frac{\alpha}{n-j+1}$ do

$j = j+1$

<u>end</u>

Reject $H_{(1)} \dots H_{(j-1)}$

Step-up procedures are more liberal than step-down procedures. <u>But</u> <u>Hochberg</u> procedure can be applied in the case when p_i <u>are independent</u>

FDR (False Discovery Rate)

	H_0 accepted	H_0 rejected	Total
H_0 true	U	V	n_0
H_0 false	T	S	$n-n_0$
	$n-R$	R	n

FDP (False discovery proportion):

$$FDP = \frac{V}{R}\mathbb{1}_{\{R \geq 1\}} = \begin{cases} \frac{V}{R}, & \text{if } R \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

FDP is unobserved random variable - proportion of false discoveries among all discoveries

FDR (False discovery rate):

$$FDR = E(FDP) = E\left(\frac{V}{R}\mathbb{1}_{\{R \geq 1\}}\right)$$

Controlling the FDR allows us to control $\frac{V}{R}$ across many repetitions of an experiment, but it does not say as much as FWER about a single experiment.

Let us show that FDR is a weaker notion of control than FWER.

Properties of the FDR:

1. Under the global null, the FDR is equivalent to the FWER.

⌈ When the global null is true then $S=0 \Rightarrow V=R$

$$\Rightarrow FDP = \begin{cases} 1, & \text{if } V \geq 1 \\ 0, & \text{otherwise} \end{cases} = \mathbb{1}_{\{V \geq 1\}} \Rightarrow$$

$\Rightarrow$

$\Rightarrow \; FDR = E(FDP) = E\mathbb{1}_{\{V \geq 1\}} = P(V \geq 1) \rfloor$

<u>Conclusion</u>: FDR control $\Rightarrow$ weak FWER control.

2. $FDR \leq FWER$.

$\ulcorner$ a) if $R = 0 \Rightarrow V = 0 \Rightarrow FDR = FWER = 0$

b) if $R \geq 1 \Rightarrow / V \leq R \Rightarrow \frac{V}{R} \leq 1 / \Rightarrow$

$\Rightarrow \frac{V}{R} \leq \mathbb{1}_{\{V \geq 1\}}$ $\left(\begin{array}{l}\text{if } V = 0 \Rightarrow 0 = 0 \\ \text{if } V \geq 1 \Rightarrow \frac{V}{R} \leq 1\end{array}\right)$

$\Rightarrow E\frac{V}{R} \leq E\mathbb{1}_{\{V \geq 1\}} = P(V \geq 1) \rfloor$

<u>Conclusion</u>: FWER control $\Rightarrow$ FDR control

<u>Benjamini - Hochberg procedure (BH)</u>

$BH(\alpha)$ - Benjamini - Hochberg procedure at level α

Let us order p-values:

$p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(n)}$

and fix $\alpha \in [0, 1]$.

Let i_0 be the largest i such that $p_{(i)} \leq \frac{i}{n}\alpha$

$i_0 = \max\{i : p_{(i)} \leq \frac{i}{n}\alpha\}$.

$\Rightarrow$ Reject $H_{(i)}$: $i \leq i_0$.

<u>Theorem 1</u>: For independent test statistics ($\Leftrightarrow p_1 \ldots p_n$ are mutually independent) $BH(\alpha)$ controls FDR at level α:

$$FDR = \frac{n_0}{n}\alpha \leq \alpha.$$

<u>Remarks</u>:

1. Theorem 1 holds for <u>all</u> configurations of the hypotheses

2. $BH(\alpha)$ threshold is adaptive.
3. Under the global null $BH(\alpha)$ controls FWER (by theorem 1). On the other hand, $BH(\alpha)$ fails to control FWER in a strong sense. (see the example for Simes procedure)
4. Comparison to Hochberg's procedure:

Hochberg's:

$p_{(1)} \le \ldots \le p_{(n)}$: $i_0 = \max\left(i : p_{(i)} \le \frac{\alpha}{n-i+1}\right)$

$BH(\alpha)$: $i_0 = \max\left(i : p_{(i)} \le \frac{i}{n}\alpha\right)$

$$\frac{i/n}{1/(n-i+1)} = i\,\frac{n-i+1}{n} = i\left(1 - \frac{i-1}{n}\right) \approx \begin{cases} i, & i \text{ is small} \\ \frac{n}{4}, & i \approx \frac{n}{2} \end{cases}$$

$\Rightarrow$ $BH(\alpha)$ is approximately i times more liberal than Hochberg's procedure when i is small and about $\frac{n}{4}$ times larger when $i \approx \frac{n}{2}$.

$\Rightarrow$ big difference from the viewpoint of power

<u>Proof of the theorem 1</u>

1) if $n_0 = 0$ (all H_{0i} are not true) $\Rightarrow V = 0$

$\Rightarrow FDR = E\,\frac{V}{R}\,\mathbb{1}_{\{R \ge 1\}} = 0 = \frac{n_0}{n}\alpha$.

2) $n_0 \ge 1$.

$V_i = \mathbb{1}_{\{H_{i0} \text{ is rejected}\}}$

$\Rightarrow FDP = \frac{V}{R}\,\mathbb{1}_{\{R \ge 1\}} = \sum_{i \in \mathcal{H}_0} \frac{V_i}{R}\,\mathbb{1}_{\{R \ge 1\}}$

$\mathcal{H}_0 = \{i : H_{0i} \text{ is true}\}$

For all $i \in \mathcal{H}_0$: $\frac{V_i}{R}\,\mathbb{1}_{\{R \ge 1\}}$ are identically distributed since p-values under H_0 have the same (uniform) distribution.

$\Rightarrow \quad E(FDP) = \sum_{i \in \mathcal{H}_0} E \frac{V_i}{R} \mathbb{1}_{\{R \geq 1\}} = n_0 \, E \frac{V_i}{R} \mathbb{1}_{\{R \geq 1\}}$

$\forall i \in \mathcal{H}_0$

$\Rightarrow$ it is enough to show that

$$E \frac{V_i}{R} \mathbb{1}_{\{R \geq 1\}} = \frac{\alpha}{n} \qquad (\forall i \in \mathcal{H}_0)$$

$\ulcorner \quad \frac{V_i}{R} \mathbb{1}_{\{R \geq 1\}} = \sum_{k=1}^{n} \frac{V_i}{k} \mathbb{1}_{\{R = k\}}$ – law of total probability

1) $R = k$ (we have k rejections) $\Rightarrow$

$\Rightarrow \quad V_i = \mathbb{1}_{\{H_{i0} \text{ is rejected}\}} = \mathbb{1}_{\{p_i \leq \frac{k\alpha}{n}\}}$.

$p_i \leq p_{(k)} \leq \frac{k\alpha}{n}$

2) let $R(p_i \to 0)$ be the number of rejections we would get from $BH(\alpha)$ if we changed p_i to 0 (keeping all the rest the same).

$p_1 \ldots p_{i-1}, \underset{\downarrow \atop 0}{p_i}, p_{i+1}, \ldots p_n$

Then $\sum_{k=1}^{n} \frac{V_i}{k} \mathbb{1}_{\{R = k\}} = \sum_{k=1}^{n} \frac{V_i}{k} \mathbb{1}_{\{R(p_i \to 0) = k\}}$

$\ulcorner$ if $V_i = 0 \Rightarrow$ true

if $V_i = 1 \Rightarrow p_i \leq \frac{k\alpha}{n} \Rightarrow p_i \to 0$ does not increase the number of rejections since p_i already below the rejection threshold whenever $V_i = 1$ $\lrcorner$

Consider $\mathcal{F}_i = \sigma\{p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n\}$ – σ-algebra generated all p-values besides p_i.

We have

$$E\left(\frac{V_i}{R}\mathbb{1}_{\{R\geq 1\}}/\mathcal{F}_i\right)=\sum_{k=1}^{n}E\left(\frac{V_i}{k}\mathbb{1}_{\{R=k\}}/\mathcal{F}_i\right)=$$

$$=\sum_{k=1}^{n}E\left(\frac{V_i}{k}\mathbb{1}_{\{R(p_i\to 0)=k\}}/\mathcal{F}_i\right)=\left|\begin{array}{l}V_i=1\Leftrightarrow\\ \Leftrightarrow p_i\leq\frac{k\alpha}{n}\end{array}\right|$$

$$=\sum_{k=1}^{n}\frac{1}{k}E\left(\mathbb{1}_{\{p_i\leq\frac{k\alpha}{n}\}}\cdot\underbrace{\mathbb{1}_{\{R(p_i\to 0)=k\}}}_{\mathcal{F}_i\text{ - measurable}}/\mathcal{F}_i\right)=$$

$$=\sum_{k=1}^{n}\frac{1}{k}\mathbb{1}_{\{R(p_i\to 0)=k\}}\cdot E\left(\mathbb{1}_{\{p_i\leq\frac{k\alpha}{n}\}}/\mathcal{F}_i\right)=$$

! p-values are independent $\Rightarrow$ does not depend on $\mathcal{F}_i$ $\Rightarrow$ $=E\,\mathbb{1}_{\{p_i\leq\frac{k\alpha}{n}\}}=P\{p_i\leq\frac{k\alpha}{n}\}$

$$=\sum_{k=1}^{n}\frac{1}{k}\mathbb{1}_{\{R(p_i\to 0)=k\}}\underbrace{P\{p_i\leq\frac{k\alpha}{n}\}}_{=\frac{k\alpha}{n}}=$$

$$=\frac{\alpha}{n}\sum_{k=1}^{n}\mathbb{1}_{\{R(p_i\to 0)=k\}}=\frac{\alpha}{n}\quad\lrcorner$$

$\lrcorner$

Proof of the last part without using the notion of $E(\ /\mathcal{F})$.

$\forall\ z_1\ldots z_{i-1}, z_{i+1},\ldots, z_n\in[0,1]$

$$E\left(\frac{V_i}{R}\mathbb{1}_{\{R\geq 1\}}\cdot\prod_{j\neq i}\mathbb{1}_{\{p_j\leq z_j\}}\right)=$$

$$=\sum_{k=1}^{n}E\left(\frac{V_i}{k}\mathbb{1}_{\{R=k\}}\prod_{j\neq i}\mathbb{1}_{\{p_j\leq z_j\}}\right)=$$

$$=\sum_{k=1}^{n}E\left(\frac{V_i}{k}\mathbb{1}_{\{R(p_i\to 0)=k\}}\prod_{j\neq i}\mathbb{1}_{\{p_j\leq z_j\}}\right)=\left|\begin{array}{l}V_i=1\Leftrightarrow\\ \Leftrightarrow p_i\leq\frac{k\alpha}{n}\end{array}\right|$$

$$=\sum_{k=1}^{n}\frac{1}{k}E\,\mathbb{1}_{\{p_i\leq\frac{k\alpha}{n}\}}\mathbb{1}_{\{R(p_i\to 0)=k\}}\prod_{j\neq i}\mathbb{1}_{\{p_j\leq z_j\}}=$$

$\Rightarrow$ $p_1 \dots p_n$ are mutually independent $\Rightarrow$

$\Rightarrow$ $\{p_i \le \frac{k\alpha}{n}\}$ and $\{R\{p_i \to 0\} = k\}, \{p_j \le z_j\}_{j \ne i}$

are independent

$$= \sum_{k=1}^{n} \frac{1}{k} \underbrace{E \mathbb{1}_{\{p_i \le \frac{k\alpha}{n}\}}}_{P\{p_i \le \frac{k\alpha}{n}\} = \frac{k\alpha}{n}} E \mathbb{1}_{\{R(p_i \to 0) = k\}} \prod_{j \ne i} \mathbb{1}_{\{p_j \le z_j\}} =$$

$$= \frac{\alpha}{n} \sum_{k=1}^{n} E \mathbb{1}_{\{R(p_i \to 0) = k\}} \prod_{j \ne i} \mathbb{1}_{\{p_j \le z_j\}} =$$ Law of Total Probability

$$= \frac{\alpha}{n} E \prod_{j \ne i} \mathbb{1}_{\{p_j \le z_j\}}$$

$$\Rightarrow E \frac{V_i}{R} \mathbb{1}_{\{R \ge 1\}} \prod_{j \ne i} \mathbb{1}_{\{p_j \le z_j\}} = \frac{\alpha}{n} E \prod_{j \ne i} \mathbb{1}_{\{p_j \le z_j\}}$$

for all $z_1 \dots z_{i-1}, z_{i+1} \dots z_n \in [0,1]$

$$\Rightarrow E \frac{V_i}{R} \mathbb{1}_{\{R \ge 1\}} = \frac{\alpha}{n}. \quad \lrcorner$$

<u>Remark</u>: In the proof we only used that $\{p_i : i \in \mathcal{H}_0\}$ are mutually independent and that $\{p_i ; i \in \mathcal{H}_0\}$ and $\{p_j : j \notin \mathcal{H}_0\}$ are independent