Lecture 38: False discovery rate (FDR)

Ciaran Evans

Outcomes for multiple hypothesis tests

Test m hypotheses, m_0 are truly null

	H_0 is true	H_0 is false	
Reject	V		R
Fail to reject			
	m_0		m

R = total # rejections

$$\text{FWER} = P(V > 0)$$

$$\text{FDP} = \begin{cases} \frac{V}{R} & \text{if } R > 0 \\ 0 & \text{if } R = 0 \end{cases}$$

(false discovery proportion)

(false discovery rate)

$$\text{FDR} = \mathbb{E}[\text{FDP}]$$

False discovery rate

(controlling FWER $\Rightarrow$ controlling FDR
controlling FDR $\not\Rightarrow$ controlling FWER)

Suppose we test m hypotheses, m_0 of which are truly null. Let V denote the number of type I errors, and R the total number of rejections.

$$FWER = P(V > 0) \qquad FDR = \mathbb{E}[FDP]$$

① If $m_0 = m$, then $FWER = FDR$

Pf: Since $m_0 = m$, either $R = 0$, or $R > 0$ and $V = R$

$$\Rightarrow \quad FDP = \begin{cases} 1 & R > 0 \\ 0 & R = 0 \end{cases}$$

$$\underbrace{\mathbb{E}[FDP]}_{FDR} = 1\,P(R>0) + 0\,P(R=0) = P(R>0) = P(V>0) = FWER \quad //$$

② In general, $FDR \leq FWER$

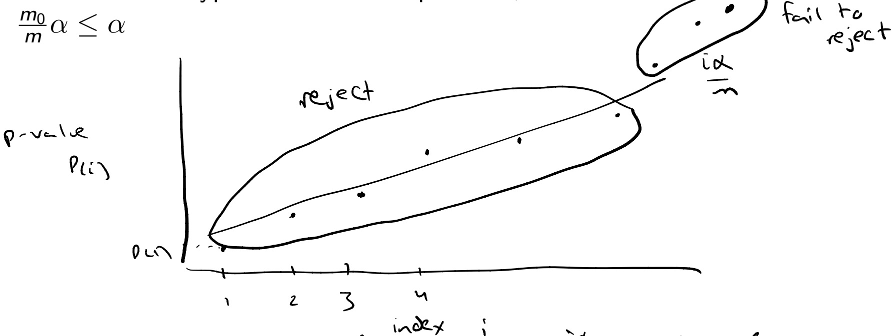
Pf: $FDP \leq \mathbb{1}\{V > 0\} \Rightarrow FDR = \mathbb{E}[FDP] \leq \mathbb{E}[\mathbb{1}\{V>0\}] = P(V>0) = FWER$

The Benjamini-Hochberg procedure

Suppose we test m null hypotheses $H_{0,1}, \ldots, H_{0,m}$. Let p_i be the corresponding p-value for test i.

- Order the p-values $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$
- Let $i^* = \max\left\{i : p_{(i)} < \frac{i\alpha}{m}\right\}$
- Reject $H_{0,(i)}$ for all $i \leq i^*$

Claim: If the hypotheses are independent, BH controls FDR at level $\frac{m_0}{m}\alpha \leq \alpha$



In this example, $p_{(6)} < \frac{6 \cdot \alpha}{m}$, $p_{(i)} > \frac{i\alpha}{m}$ $\forall\, i > 6$

Intuition: If H_0 is true, we expect p-values to be $U(0,1)$

$$i^* = \max\left\{i : P_{(i)} < \frac{i\alpha}{m}\right\}$$ => reject all p-values P_j

$$P_j < \frac{i^*\alpha}{m}$$

If H_0 is true, $$P\left(p_j < \frac{i^*\alpha}{m}\right) = \frac{i^*\alpha}{m}$$

we have m_0 true null hypotheses

=> we expect $m_0 \cdot \frac{i^*\alpha}{m}$ type I errors

$$\Rightarrow \text{FDP} \approx \frac{m_0 i^* \alpha / m}{i^*}$$ $i^* = \#$ rejections total

$$\approx \frac{m_0 \alpha}{m} \leq \alpha$$

$$\text{FDR} \leq \frac{m_0 \alpha}{m} \leq \alpha$$

Summary

- BH controls FDR at level $\frac{m_0}{m}\alpha$
- If $m_0 = m$, then controlling FDR is equivalent to controlling FWER
- If $m_0 < m$, then controlling FDR provides more power to reject H_0 when H_0 is false