April 18, 2019

1 False Discovery Rate

Testing error

Hypothesis	*Fail to Reject*	*Reject*	*Total*
Null true	U	V	m_0
Alternative true	T	S	m_1
	W	R	m

Figure 1: Testing Error

Q is The proportion of the rejected null hypothesis which are erroneously rejected

$$Q = \frac{V}{R} = \frac{V}{V+S}\ ,\ FDR = E[Q] = E[\frac{V}{R}]$$

- Useful approach to simultaneous testing
- relies on p-values that is on null hypothesis areas

2 Local FDR

$$\text{Null Hypothesis} : H_1, H2, \dots, H_i, \dots, H_N$$

$$\text{Test Statistics} : z_1, z_2, \dots, z_i, \dots, z_N$$

We assume that th N cases are divided into two classes, null or non-null, occurring with prior probabilities p_0 or $p_1 = 1 - p_0$ and density of test statistic z depending upon its class.

$$p_0 = Pr\{null\}\ f_0(z)\text{density if null}$$

$$p_1 = Pr\{non - null\}\ f_1(z)\text{density if non-null}$$

We can write the null subdensity as,

$$f_0^+(z) = p_0 f_0(z)$$

and the mixture density

$$f(z) = p_0 f_0(z) + p_1 f_1(z)$$

by definition the local false discovery rate is,

$$\begin{aligned} fdr(z) &\equiv Pr\{null|z\} = p_0 f_0(z)/f(z) \\ &= f_0^+(z)/f(z) \end{aligned}$$

Letting the $F_0(z)$ and $F_1(z)$ be the cdf's define $F_0^+ = p_0 F_0(z)$ and $F(z) = p_0 F_0(z) + p_1 F_1(z)$

$$\begin{aligned} Fdr(z) &\equiv Pr\{null|Z \leq z\} = p_0 F_0(z)/F(z) \\ Fdr(z) &= \int_{-\infty}^{z} fdr(Z) f(Z) dZ / \int_{-\infty}^{z} f(Z) dZ \\ &= E_f\{fdr(Z)|Z \leq z\} \end{aligned}$$

$Fdr(z)$ is the average of $fdr(z)$ for $Z \leq z$, $Fdr(z)$ will be less than $fdr(z)$ in the usual situation where $fdr(z)$ decreases as $|z|$ gets large.

3 Example

$$\begin{aligned} Y_i|\beta_i &\overset{ind}{\sim} N(\beta_i, 1) \\ \beta_i &\sim \begin{cases} 0 & \text{with probability } 0.8 \\ N(0,1) & \text{with probability } 0.2 \end{cases} \end{aligned}$$

We can write $Y_i = \beta_i + \epsilon_i$, where $\epsilon_i \sim N(-, 1)$

- $p_0 = Pr\{null\} = 0.8$
- $p_1 = Pr\{non - null\} = 0.2$
- $f_0(y) \sim N(0, 1)$ density of Y_i if null
- $f_1(y) \sim N(0, 2)$ density of Y_i if non-null
- $f(y) = p_0 f_0(y) + p_1 f_1(y)$

Local FDR is

$$fdr(y) = p_0 f_0(y)/f(y) = Pr\{null|Y_i = y\}$$
$$= Pr(\beta_i = 0|Y) = \frac{0.8 \cdot N(0,1)}{0.8 \cdot N(0,1) + 0.2 \cdot N(0,2)}$$

For instance $fdr(0) = \frac{0.8 \cdot f_0(0)}{0.8 \cdot f_0(0) + 0.2 \cdot f_1(0)} = 0.8888889$

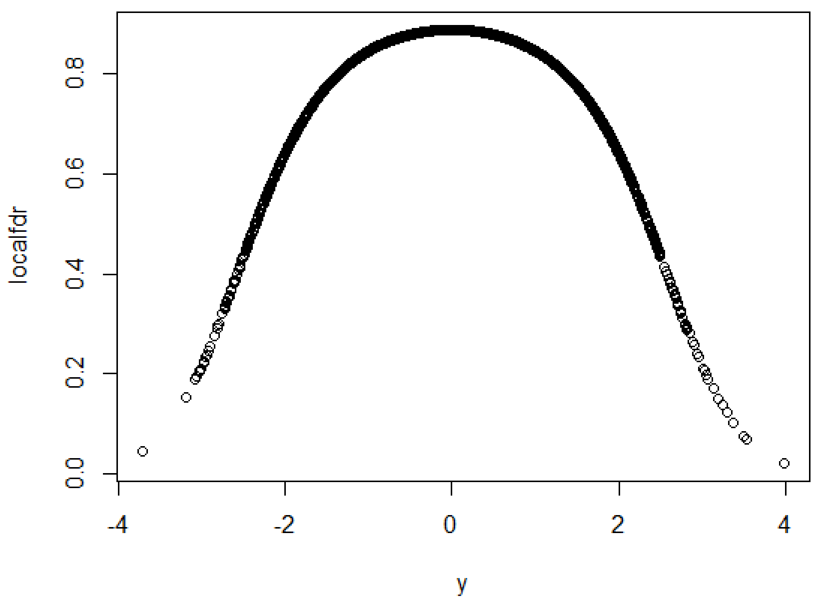

Figure 2: N = 10000