單元 24: 統計檢定的要素
(課本 §10.2)

以例說明:

1. 1 號候選人宣稱其

得票率 $\geq$ 50% (虛無假設, null hypothesis)

2. 我們不相信此宣稱, 此乃相當於此後選人的

得票率 $<$ 50%(研究假設, research hypothesis)

或稱對立假設 (alternative hypothesis), 這是重點, 關注的事.

3. 隨機選取 15 位合格選民並得

$$Y = \text{樣本中支持 1 號的選民數}$$

稱作檢定統計量 (test statistic).

4. 判斷依據爲, 若 $Y = 0$, 則拒絕 1 號的宣稱而支持對立假設 (reject null hypothesis in favor of alternative hypothesis);

若 $Y = 1$, 則拒絕 1 號的宣稱而支持對立假設 (reject null hypothesis in favor of alternative hypothesis);

若 $Y = 2$, 則拒絕 1 號的宣稱而支持對立假設 (reject null hypothesis in favor of alternative hypothesis).

也就是說, 設定棄絕域 (rejection region, RR, 拒絕域).

綜合, 歸納出統計檢定的四要素爲

(1) 虛無假設 (null hypothesis) H_0 ($H_0 : p = 0.5$)

(2) 對立假設 (alternative hypothesis, 又稱 research hypothesis) H_a ($H_a : p < 0.5$)

(3) 檢定統計量 (test statistic) (Y, 樣本觀察值)

(4) 拒絕域 (rejection region, 棄絕域) RR (RR $= \{y \leq 2\}$, 用以拒絕 H_0 的觀察值範圍)

問. 如何選取一個好的拒絕域 RR? 此乃相當於如何降低下述二種錯誤決策的機率?

定義. 型 I 錯誤 (type I error): 拒絕 H_0, 當 H_0 爲眞時;

$$P(\text{型 I 錯誤}) = \alpha$$

又稱作檢定水準 (level of test) 或意義水準 (level of significance).

型 II 錯誤 (type II error): 接受 H_0, 當 H_a 爲眞時;

$$P(\text{型 II 錯誤}) = \beta$$

例 1. 民調 $n = 15$ 位選民且檢定

$$H_0 : p = 0.5 \text{ 對立於 } H_a : p < 0.5$$

其中 p 爲 1 號的得票率. 則支持 1 號的選民數

$$Y \sim \text{binomial}(15, p)$$

(a) 試計算 α 若 RR $= \{y \leq 2\}$.

(b) 試求 β 若 RR $= \{y \leq 2\}$ 且 1 號的眞實得票率爲 30%, 即 $p = 0.3$.

<解> (a) 由定義,

$$\begin{aligned}
\alpha &= P(\text{型 I 錯誤}) = P(Y \in \text{RR},\ \text{當 } H_0 \text{ 眞時}) \\
&= P(Y \leq 2,\ \text{當 } p = 0.5) \\
&\quad (\text{故 } Y \sim \text{binomial}(15, 0.5)) \\
&= \sum_{y=0}^{2} \binom{15}{y} (0.5)^y (0.5)^{15-y} = 0.004
\end{aligned}$$

由表 1, 或

R: `pbinom(2,15,0.5)`

此乃說明在當選下而作出落選的判斷的風險是非常的小.

(b) 由定義,

$$\begin{aligned}
\beta &= P(\text{型 II 錯誤}) = P(Y \notin \text{RR},\ \text{當 } H_a \text{ 爲眞}) \\
&= P(Y > 2,\ \text{當 } p = 0.3) \\
&\quad (\text{故 } Y \sim \text{binomial}(15, 0.3)) \\
&= \sum_{y=3}^{15} \binom{15}{y} (0.3)^y (0.7)^{15-y} \\
&= 1 - \sum_{y=0}^{2} \binom{15}{y} (0.3)^y (0.7)^{15-y} \\
&= 1 - 0.127 = 0.873
\end{aligned}$$

由表 1, 或

R: `pbinom(2, 15, 0.3, lowertail=F)`

此乃表示有相當大的風險作出即使支持率 30% 而判斷當選的錯誤決策.

(c) 試求 β 若 RR $= \{y \leq 2\}$ 且 1 號的眞實支持率爲 10%, 即 $p = 0.1$.

(d) 試求 α^* 與 β^* 若 RR$^* = \{y \leq 5\}$ 且 1 號的眞實得票率爲 30%, 即 $p = 0.3$.

<解> 此時

$$\begin{aligned}
\beta &= P(Y > 2,\ \text{當}\ p = 0.1) \\
&\quad (\text{故}\ Y \sim \text{binomial}(15, 0.1)) \\
&= 1 - \sum_{y=0}^{2} \binom{15}{y} (0.1)^y (0.9)^{15-y} \\
&= 1 - 0.816 = 0.184
\end{aligned}$$

由表 1, 或

R: `pbinom(2, 15, 0.1, lowertail=F)`

依然相當地可能作出當選的錯誤結論即使 $p = 0.1$.

註. 型 I 錯誤的機率 $\alpha = 0.004$ 相當小, 好; 但型 II 錯誤的機率

$$\beta = \begin{cases} 0.873, & \text{若}\ p = 0.30 \\ 0.184, & \text{若}\ p = 0.10 \end{cases}$$

均過大, 不好.

問. 如何改進?

答. 設法改變 RR 而平衡 α 與 β, 如下述.

(1) 選取一新的棄絕域 $RR^* \supset$ RR, 則

$$\begin{aligned} \alpha^* &= P(\text{檢定統計量} \in \text{RR}^*,\ \text{當 } H_0 \text{ 真時}) \\ &\geq P(\text{檢定統計量} \in \text{RR},\ \text{當 } H_0 \text{ 真時}) = \alpha \end{aligned}$$

且

$$\begin{aligned} \beta^* &= P(\text{檢定統計量} \notin \text{RR}^*,\ \text{當 } H_a \text{ 真時}) \\ &\leq P(\text{檢定統計量} \notin \text{RR},\ \text{當 } H_a \text{ 真時}) = \beta \end{aligned}$$

如圖示.

(2) 同理, 若選取 RR$^* \subset$ RR, 則

$$\alpha^* < \alpha \text{ 且 } \beta < \beta^*$$

如圖示.

結論爲, 不可能同時使得 α^* 與 β^* 均小.

(d) 首先,

$$\begin{aligned}\alpha^* &= P(Y \le 5, \text{ 當 } p = 0.5) \\ &= \sum_{y=0}^{5} \binom{15}{y} (0.5)^y (0.5)^{15-y} \\ &= 0.151 > 0.004 = \alpha\end{aligned}$$

接著,

$$\begin{aligned}\beta^* &= P(Y > 5, \text{ 當 } p = 0.3) \\ &= 1 - \sum_{y=0}^{5} \binom{15}{y} (0.3)^y (0.7)^{15-y} \\ &= 1 - 0.722 = 0.278 < 0.873 = \beta\end{aligned}$$

如預期且 α^* 與 β^* 均太大, 不好.

問. 如何同時降低 α 與 β?

答. 一個直覺的回答是, 增加樣本大小以闡明更多母體的本質而達到同時降低犯錯率的目的.