單元 25: 常見的大樣本檢定
(課本 §10.3)

設隨機變數

$$Y_1, \ldots, Y_n$$

爲一由含參數 θ 的母體所得的隨機樣本. 考慮 θ 的不偏估計量 $\widehat{\theta}$ 且滿足當 $n \to \infty$,

$$\frac{\widehat{\theta} - \theta}{\sigma_{\widehat{\theta}}} \xrightarrow{\mathcal{D}} N(0, 1)$$

意即, 當 n 夠大時, $\widehat{\theta}$ 的 pdf $f(\widehat{\theta})$ 如圖示. 例如,

1. $\widehat{\theta} = \widehat{\mu} = \overline{Y}$ (樣本期望值), $\sigma_{\widehat{\theta}} = \sigma_{\widehat{\mu}} = \frac{\sigma}{\sqrt{n}}$

2. $\widehat{\theta} = \widehat{p}$ (樣本比率, sample proportion),
$\sigma_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$

3. $\widehat{\theta} = \widehat{\mu}_1 - \widehat{\mu}_2$, $\sigma_{\widehat{\theta}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

4. $\widehat{\theta} = \widehat{p}_1 - \widehat{p}_2$, $\sigma_{\widehat{\theta}} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

一假設的大樣本檢定 (又稱作 Z 檢定). 令 θ_0 爲 θ 的某一特定值.

一. 檢定

$$H_0 : \theta = \theta_0 \text{ 對立於 } H_a : \theta > \theta_0$$

稱作上尾檢定 (upper-tail test).

檢定統計量: $\widehat{\theta}$, 不偏並滿足當 $n \to \infty$ 時,

$$\frac{\widehat{\theta} - \theta}{\sigma_{\widehat{\theta}}} \xrightarrow{\mathcal{D}} N(0,1)$$

拒絕域: RR $= \{\widehat{\theta} > k\}$, k 爲某一大於 θ_0 的常數.

爲何是 $\widehat{\theta} > k$ 的型式? 直覺上, 如 pdf $f(\widehat{\theta})$ 的圖示,

(1) 當 $\theta \approx \theta_0$ 時, $\widehat{\theta} \approx \theta_0$, 得較支持 H_0, 故接受 H_0.

(2) 當 $\theta \gg \theta_0$, $\widehat{\theta} \gg \theta_0$, 得缺乏 H_0 的支持, 故拒絕 H_0, 而得 RR 的型式爲 $\widehat{\theta} > k$, k 爲某一大於 θ_0 的常數.

問. 如何選取一個好的 k?

答. 根據給定的 $\alpha = P$(型 I 錯誤) 選定如下. 首先, 由定義, 當 n 夠大時,

$$\begin{aligned}
\alpha &= P(\widehat{\theta} \in \text{RR}, \text{ 當 } H_0 \text{ 眞時}) \\
&= P(\widehat{\theta} > k, \text{ 當 } \theta = \theta_0) \\
&= P\left(\frac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}} > \frac{k - \theta_0}{\sigma_{\widehat{\theta}}}, \text{ 當 } \theta = \theta_0\right) \\
&\approx P\left(N(0,1) > \frac{k - \theta_0}{\sigma_{\widehat{\theta}}}\right)
\end{aligned}$$

接著, 若 z_α 滿足

$$P(N(0,1) > z_\alpha) = \alpha$$

則由上式得

$$\frac{k - \theta_0}{\sigma_{\widehat{\theta}}} \approx z_\alpha$$

即

$$k \approx \theta_0 + z_\alpha \sigma_{\widehat{\theta}}$$

因此, 可選 $k = \theta_0 + z_\alpha \sigma_{\widehat{\theta}}$ 以保證一小的

$$P(\text{型 I 錯誤}) \approx \alpha \text{ (給定的值)}$$

且

$$\begin{aligned}\text{RR} &= \{\widehat{\theta}:\widehat{\theta}>k\}=\{\widehat{\theta}:\widehat{\theta}>\theta_0+z_\alpha\sigma_{\widehat{\theta}}\}\\ &= \left\{\widehat{\theta}:\frac{\widehat{\theta}-\theta_0}{\sigma_{\widehat{\theta}}}>z_\alpha\right\}\end{aligned}$$

最後, 在 $\alpha=P$(型 I 錯誤) 下, 一等價的大樣本假設檢定爲

$H_0:\theta=\theta_0$

$H_a:\theta>\theta_0$ (上尾對立假設)

檢定統計量: $Z\stackrel{\text{def}}{=}\dfrac{\widehat{\theta}-\theta_0}{\sigma_{\widehat{\theta}}}$, 即 $\widehat{\theta}$ 與 θ_0 間的標準差數.

拒絕域: RR $=\{z>z_\alpha\}$, 其中 z_α 滿足

$$P(N(0,1)>z_\alpha)=\alpha$$

如圖示.

例 1. 某公司副總裁宣稱每位業務員的每週業績小於或等於 15 件. 他希望能增加此件數. 經由取 $n=36$ 的樣本, 得

$$\overline{y}=17 \text{ 且 } s^2=9$$

試問在選取 $\alpha = 0.05$ 下, 樣本證據是否足以反駁副總裁的宣稱?

<解> 由題意, 檢定

$$H_0 : \mu = 15 \text{ 對立於 } H_a : \mu > 15$$

因爲 $n = 36$ 是夠大, 故 $\overline{Y}$ 滿足

$$\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \xrightarrow{\mathcal{D}} N(0,1)$$

因此, 可考慮大樣本檢定並取

$$Z = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}}$$

作爲檢定統計量. 又因爲

$$P(N(0,1) > 1.645) = 0.05 = \alpha$$

得

$$\text{RR} = \{z > 1.645\}$$

最後, 因爲 σ 未知, 在夠大的 $n = 36$ 下, 可以 s 取代, 而得 Z 的觀察値

$$\begin{aligned} z &= \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}} \\ &\approx \frac{\overline{y} - \mu_0}{s/\sqrt{n}} = \frac{17 - 15}{\sqrt{9/36}} = 4 \in \text{RR} \end{aligned}$$

故資料顯示充分證據拒絕 H_0, 即反駁副總裁的宣稱.

例 2. 若瑕疵率超過 10%, 則機器送修. 今隨機抽取 100 件產品, 得 15 件有瑕疵且督導宣稱機器需送修. 試問在 $\alpha = 0.01$ 下, 樣本證據是否支持督導的決策?

<解> 由題意, 檢定

$$H_0 : p = 0.1 \text{ 對立於 } H_a : p > 0.1$$

因爲 $n = 100$ 夠大, 故不偏估計量 $\widehat{p}$ 滿足

$$\frac{\widehat{p} - p}{\sigma_{\widehat{p}}} = \frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{\mathcal{D}} N(0,1)$$

因此, 可考慮大樣本檢定並取

$$Z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

作爲檢定統計量. 又查表得

$$P(N(0,1) > 2.33) = 0.01 = \alpha$$

或

R: $z_{0.01}$ `= qnorm(.99, 0, 1)`

故

$$\text{RR} = \{z > 2.33\}$$

最後, Z 的觀察值

$$
\begin{aligned}
z &= \frac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.15-0.10}{\sqrt{(0.10)(0.90)/100}} \\
&= \frac{0.05}{0.3/10} = \frac{5}{3} \notin \text{RR}
\end{aligned}
$$

故在意義水準 $\alpha = 0.01$ 下, 無法拒絕 H_0, 意即無充分證據支持督導的決策.

二. 檢定

$$H_0 : \theta = \theta_0 \text{ 對立於 } H_a : \theta < \theta_0$$

稱作下尾檢定 (lower-tail test).

檢定統計量: $\widehat{\theta}$, 不偏並滿足當 $n \to \infty$ 時,

$$\frac{\widehat{\theta}-\theta}{\sigma_{\widehat{\theta}}} \xrightarrow{\mathcal{D}} N(0,1)$$

拒絕域: $\text{RR} = \{\widehat{\theta} < k\}$, k 爲某一小於 θ_0 的常數.

爲何是 $\widehat{\theta} < k$ 的型式? 如圖示,

(1) 當 $\theta \approx \theta_0$ 時, $\widehat{\theta} \approx \theta_0$, 得較支持 H_0, 故接受 H_0.

(2) 當 $\theta \ll \theta_0$, $\hat{\theta} \ll \theta_0$, 得缺乏 H_0 的支持, 故拒絕 H_0, 而得 RR 的型式爲 $\hat{\theta} < k$, $k < \theta_0$.

給定 $\alpha = P$(型 I 錯誤), 則由定義, 當 n 夠大時,

$$\begin{aligned}
\alpha &= P(\hat{\theta} \in \text{RR}, \text{ 當 } H_0 \text{ 眞時}) \\
&= P(\hat{\theta} < k, \text{ 當 } \theta = \theta_0) \\
&= P\left(\frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} < \frac{k - \theta_0}{\sigma_{\hat{\theta}}}, \text{ 當 } \theta = \theta_0\right) \\
&\approx P\left(N(0,1) < \frac{k - \theta_0}{\sigma_{\hat{\theta}}}\right)
\end{aligned}$$

若 z_α 滿足

$$P(N(0,1) < -z_\alpha) = \alpha$$

如圖示, 則由上式得

$$\frac{k - \theta_0}{\sigma_{\hat{\theta}}} \approx -z_\alpha$$

即

$$k \approx \theta_0 - z_\alpha \sigma_{\hat{\theta}}$$

因此, 可選取 $k = \theta_0 - z_\alpha \sigma_{\hat{\theta}}$ 以保證

$$P(\text{型 I 錯誤}) \approx \alpha \text{ (給定的值)}$$

且

$$\begin{aligned}\text{RR} &= \{\widehat{\theta} : \widehat{\theta} < k\} = \{\widehat{\theta} : \widehat{\theta} < \theta_0 - z_\alpha \sigma_{\widehat{\theta}}\} \\ &= \left\{\widehat{\theta} : \frac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}} < -z_\alpha\right\}\end{aligned}$$

最後, 在 α 下, 一等價的大樣本假設檢定為

$H_0 : \theta = \theta_0$

$H_a : \theta < \theta_0$ (下尾對立假設)

檢定統計量: $Z \stackrel{\text{def}}{=} \dfrac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}}$, 即 $\widehat{\theta}$ 與 θ_0 間的標準差數.

拒絕域: $\text{RR} = \{z < -z_\alpha\}$, 其中 z_α 滿足

$$P(N(0,1) > z_\alpha) = \alpha$$

如圖示.

三. 檢定

$$H_0 : \theta = \theta_0 \text{ 對立於 } H_a : \theta \neq \theta_0$$

稱作雙尾檢定 (two-tailed test).

檢定統計量: $\widehat{\theta}$, 不偏並滿足當 $n \to \infty$ 時,

$$\frac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}} \xrightarrow{\mathcal{D}} N(0, 1)$$

拒絕域: $\text{RR} = \{\widehat{\theta} < k_1 \text{ 或 } \widehat{\theta} > k_2\}$, 其中 $k_1 < \theta_0$ 且 $k_2 > \theta_0$.

同理, 基於 $\alpha = P(\text{型 I 錯誤})$, 得等價的大樣本假設檢定爲

$H_0 : \theta = \theta_0$

$H_a : \theta \neq \theta_0$ (雙尾對立假設)

檢定統計量: $Z \overset{\text{def}}{=} \dfrac{\widehat{\theta} - \theta_0}{\sigma_{\widehat{\theta}}}$

拒絕域: $\text{RR} = \{|z| > z_{\alpha/2}\}$, 其中 $z_{\alpha/2}$ 滿足

$$P(N(0, 1) > z_{\alpha/2}) = \alpha/2$$

如圖示.

例 3. 一心理學研究測量男性與女性對某一刺激的反應時間, 獨立地抽樣資料如下,

男性	女性
$n_1 = 50$	$n_2 = 50$
$\overline{y}_1 = 3.6$	$\overline{y}_2 = 3.8$
$s_1^2 = 0.18$	$s_2^2 = 0.14$

試問在 $\alpha = 0.05$ 下, 男性與女性的平均反應時間有差異嗎?

<解> 由題意, 檢定

$$H_0 : \mu_1 = \mu_2 \text{ 對立於 } H_a : \mu_1 \neq \mu_2$$

令 $d = \mu_1 - \mu_2$, 此乃相當於檢定

$$H_0 : d = 0 \stackrel{\text{def}}{=} d_0 \text{ 對立於 } H_a : d \neq 0 = d_0$$

因爲 $n_1 = n_2 = 50$ 夠大, 故不偏估計量 $\overline{Y}_1 - \overline{Y}_2$ 滿足

$$\frac{(\overline{Y}_1 - \overline{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \xrightarrow{\mathcal{D}} N(0,1)$$

因此, 可考慮大樣本檢定並取

$$Z = \frac{(\overline{Y}_1 - \overline{Y}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

作爲檢定統計量. 又查表得

$$P(N(0,1) > 1.96) = 0.025 = \alpha/2$$

或

$$\text{R: } z_{0.025} = \texttt{qnorm(0.975, 0, 1)}$$

故

$$\text{RR} = \{|z| > 1.96\}$$

最後, 由樣本資料, 在夠大的 $n_1 = n_2 = 50 > 30$ 下, 以 s_1 與 s_2 取代未知的 σ_1 與 σ_2, 得 Z 的觀察值

$$\begin{aligned} z &= \frac{(\bar{y}_1 - \bar{y}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \\ &\approx \frac{(\bar{y}_1 - \bar{y}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{3.6 - 3.8 - 0}{\sqrt{\frac{.18}{50} + \frac{.14}{50}}} \\ &= -2.5 < -1.96 \end{aligned}$$

故 $z \in \text{RR}$ 且在意義水準 $\alpha = 0.05$ 下拒絕 H_0, 意即有充分證據支持男性與女性的平均反應時間有差異.