單元 32: 檢定力與 Neyman-Pearson 引理
(課本 §10.10)

針對未知參數 θ 的統計檢定程序的四要素為

1. H_0, 虛無假設, 如 $H_0 : \theta = \theta_0$.

2. H_a, 對立假設, 如 $H_a : \theta = \theta_a$.

3. 檢定統計量 W, 一 $Y_1, \ldots, Y_n$ 的函數.

4. 拒絕域 RR, 若 $W \in$ RR, 則拒絕 RR; 由 α 決定.

一檢定的優良性 (goodness) 可由二錯誤機率判斷, 即

(1) 型 I 錯誤機率

$$\begin{aligned} \alpha &= P(\text{拒絕 } H_0,\ \text{當 } H_0 \text{ 真時}) \\ &= P(W \in \text{RR},\ \text{當 } H_0 \text{ 真時}) \end{aligned}$$

(2) 型 II 錯誤機率

$$\begin{aligned}\beta &= P(\text{接受 } H_0, \text{ 當 } \theta = \theta_a, \\ &\qquad\qquad \text{某 } H_a \text{ 中的特定值 } \theta_a) \\ &= P(W \notin \text{RR}, \text{ 當 } \theta = \theta_a), \text{ 因 } \theta_a \text{ 而定}\end{aligned}$$

註 1. 通常事先選定 α 並根據此 α 決定 RR 且當樣本大小固定時,

$$\alpha \uparrow, \beta \downarrow \text{ 或 } \alpha \downarrow, \beta \uparrow$$

即無法同時降低二種錯誤機率, 如圖示, 針對上尾檢定的說明.

註 2. 可選取更大的樣本大小 n, 以便同時降低 α 與 β, 即透過增大 n 使變異數降低的效應, 讓 W 更集中更陡峭, 而得小的 α 與 β, 如圖示.

另一更有用的評估檢定的方法稱作一檢定的檢定力 (power of a test), 定義如下.

定義. 檢定力

$$\text{power}(\theta) \stackrel{\text{def}}{=} P(W \in \text{RR}, \text{ 當參數值是 } \theta \text{ 時})$$

即眞實參數值爲 θ 時, 拒絕 H_0 的機率, 乃一 θ 的函數.

註 1. 由定義,

$$\text{power}(\theta_0) = P(W \in \text{RR},\ \text{當}\ \theta = \theta_0) = \alpha$$

註 2. 針對 $\theta_a \in H_a$,

$$\begin{aligned}\text{power}(\theta_a) &= P(W \in \text{RR},\ \text{當}\ \theta = \theta_a) \\ &= 1 - P(W \notin \text{RR},\ \text{當}\ \theta = \theta_a) \\ &= 1 - \beta\end{aligned}$$

即一檢定能偵測出虛無假設是錯誤的能力.

註 3. 檢定

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

時, 典型的檢定力曲線與理想的檢定力如圖示.

問. 針對固定的樣本大小, 如何求一優良的檢定? 一般而言, 構想爲

(1) 事先選定一小的 α (意義水準).

(2) 針對 H_a 中的 θ_a, 求 RR 以最小化 $\beta(\theta_a)$, 亦相當於針對 H_a 中的 θ_a, 求 RR 以最大化 power(θ_a).

在落實上述二構想前, 需先定義下列各項.

定義. 令 $Y_1, \ldots, Y_n$ 爲一由含參數 θ 的母體所取得的隨機樣本. 一假設 H 稱作簡單的 (simple) 若 H 可唯一地確定母體的分布; 否則, 稱作複合的 (composite). 例如,

1. 設隨機變數

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{exp}(\lambda)$$

即 pdf

$$f(y) = \begin{cases} \frac{1}{\lambda} e^{-y/\lambda}, & y > 0 \\ 0, & \text{其它} \end{cases}$$

若

$$H_1 : \lambda = 2$$

則在 H_1 下, pdf

$$f(y) = \frac{1}{2} e^{-y/2}, \ y > 0$$

可唯一地確定, 故 H_1 爲一簡單假設. 若

$$H_2 : \lambda > 2$$

則在 H_2 下, pdf

$$f(y) = \frac{1}{\lambda} e^{-y/\lambda}, \ y > 0$$

某個 $\lambda > 2$, 無法唯一地確定, 故 H_2 是一個複合假設.

2. 設隨機變數

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

若 σ^2 已知且

$$H : \mu = 1$$

則 pdf

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-1)^2/2\sigma^2}$$

唯一地確定, 故 H 是簡單的. 若 σ^2 未知且

$$H : \mu = 1$$

則 pdf

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-1)^2/2\sigma^2}$$

無法確定, 故 H 是複合的.

定義. 檢定

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

即簡單對簡單時, 若一固定的拒絕域 RR 滿足

$$\alpha = \text{power}(\theta_0)$$

且針對所有意義水準爲 α 的拒絕域,

$$\text{power}(\theta_a) = P(\text{RR, 當 } \theta = \theta_a)$$

是最大的, 則以 RR 作爲拒絕域的檢定稱作 α 水準最力檢定 (most powerful α-level test), 簡記作 α-水準 MPT.

問. 如何求 MPT?

答. 可根據下述的 Neyman-Pearson 引理求得.

定理 10.1 Neyman-Pearson 引理. 假設由含 θ 的母體所取得的隨機樣本

$$Y_1, \ldots, Y_n$$

檢定

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

即簡單對簡單. 令

$$L(\theta) = L(y_1, \ldots, y_n | \theta)$$

爲參數值是 θ 時樣本的概似值 (likelihood). 則對於給定的 α, 最大化 power(θ_a) 的檢定的拒絕域

$$\text{RR} = \left\{ \frac{L(\theta_0)}{L(\theta_a)} < k \right\}$$

其中 k 滿足

$$P\left(\frac{L(\theta_0)}{L(\theta_a)} < k, \text{ 當 } \theta = \theta_0 \text{ 時} \right) = \alpha$$

並稱以此 RR 爲拒絕域的檢定爲 α-水準最力檢定 (most powerful α-level test).

例 1. 設單一觀察值

$$Y \sim f(y) = \begin{cases} \theta y^{\theta-1}, & 0 < y < 1 \\ 0, & \text{其它} \end{cases}$$

試在意義水準 α 下, 求檢定

$$H_0 : \theta = 2 \text{ 對立於 } H_a : \theta = 1$$

的 MPT.

<解> 因爲 H_0 與 H_a 均是簡單的, 故由 Neyman-Pearson 引理, $\alpha = 0.05$ 的 MPT 的

$$\begin{aligned} \text{RR} &= \left\{ 0 < y < 1 : \frac{L(\theta_0)}{L(\theta_a)} = \right. \\ &\qquad \left. \frac{f(y|\theta_0)}{f(y|\theta_a)} = \frac{2y}{1y^0} = 2y < k \right\} \end{aligned}$$

整理, 亦相當於

$$\text{RR} = \{0 < y < 1 : y < k'\}$$

其中 k' 使得

$$\begin{aligned} \alpha &= 0.05 = P(\text{RR},\ \text{當 } \theta = \theta_0) \\ &= P(Y < k',\ \text{當 } \theta = 2) \\ &= \int_0^{k'} f(y|\theta = 2) dy = \int_0^{k'} 2y dy = (k')^2 \end{aligned}$$

即

$$k' = \sqrt{0.05} = 0.2236$$

因此, MPT 的

$$\text{RR} = \{y < 0.2236\}$$

意即在所有 $\alpha = 0.05$ 的檢定中, 此 RR 的

$$\begin{aligned} \text{power}(1) &= P(Y \in \text{RR},\ \text{當 } \theta = 1) \\ &= P(Y < 0.2236,\ \text{當 } \theta = 1) \\ &= \int_0^{0.2236} 1 dy = 0.2236 \end{aligned}$$

是最大的, 但

$$\beta(1) = 1 - 0.2236 = 0.7764$$

還是太大. 當然其它的檢定會有更大的 β 值.

註 1. 檢定統計量與 RR 的型式因 H_0 與 H_a 而定, 如檢定

$$H_0 : \theta = 2 \text{ 對立於 } H_a : \theta = 4$$

時, 拒絕域

$$\begin{aligned} \text{RR} &= \left\{ 0 < y < 1 : \frac{f(y|\theta = 2)}{f(y|\theta = 4)} = \frac{2y}{4y^3} < k \right\} \\ &= \{ 0 < y < 1 : y^2 > k' \} \end{aligned}$$

其中 k' 滿足

$$\alpha = 0.05 = \int_{\sqrt{k'}}^{1} 2y dy = 1 - k'$$

故 $k' = 0.95$ 且拒絕域

$$\text{RR} = \{ y^2 > 0.95 \}$$

以及檢定統計量爲 Y^2.

註 2. 對於離散母體, 以 P(RR, 當 H_0 眞時) 不超過且最靠近事先給定的 α 值的 RR 作爲 MPT.

註 3. 檢定

$$H_0 : \theta = \theta_0 \text{ (簡單) 對立於 } H_a : \theta > \theta_0 \text{ (複合)}$$

時, 若一固定的 RR 使得 $\alpha = \text{power}(\theta_0)$ 且對每一個 $\theta_a > \theta_0$, 在所有意義水準爲 α 的檢定中,

$$\text{power}(\theta_a) = P(RR, \text{ 當 } \theta = \theta_a)$$

都是最大的, 則稱以 RR 爲拒絕域的檢定爲 α-水準一致最力檢定 (uniformly most powerful α-level test), 簡記作 UMPT. 同理, 可定義檢定

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

的 UMPT.

註 4. 假設母體除了單一參數 θ 外, 可被完全地確定, 問檢定

$$H_0 : \theta = \theta_0 \text{ (簡單) 對立於 } H_a : \theta > \theta_0 \text{ (複合)}$$

時, 如何求 UMPT?

答. 沒有如同 Neyman-Pearson 引裡的廣義定理可用, 但若對於每一個 $\theta_a > \theta_0$, 由 Neyman-Pearson 引理所導出, 用來檢定

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

的 MPT 的 RR 均有相同的型式, 即不因 θ_a 而定, 則此 RR 所對應的檢定即是一檢定

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

的 UMPT.

爲何如此? 若對所有的 $\theta_a > \theta_0$,

$$\text{RR} = \left\{ \frac{L(\theta_0)}{L(\theta_a)} < k(\theta_a) \right\}$$

即型式相同, 則由

$$\alpha = P\left(\frac{L(\theta_0)}{L(\theta_a)} < k(\theta_a), \text{ 當 } \theta = \theta_0 \text{ 時} \right)$$

可導出單一的 RR, 且此 RR 對所有的 $\theta_a > \theta_0$, 在所有意義水準爲 α 的檢定中, 都有最大的 power(θ_a), 故 RR 是 UMPT.

例 2. 設隨機變數

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

其中 μ 未知, σ^2 已知. 試求在意義水準 α 下, 檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu > \mu_0$$

的 UMPT.

註. 由過去經驗, 針對上尾檢定, 可採用檢定統計量

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

以及

$$\text{RR} = \{z > z_\alpha\}$$

其中 $P(N(0,1) > z_\alpha) = \alpha$. 此處將證明此 Z-檢定就是一 UMPT.

<解> 首先, 求

$$H_0 : \theta = \mu_0 \text{ 對立於 } H_a : \theta = \mu_a$$

某固定的 $\mu_a > \mu_0$ 的 MPT. 因爲 σ 已知, 故 Y_i 的共同 pdf

$$f(y|\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2},\ -\infty < y < \infty$$

且得

$$\begin{aligned}
L(\mu) &= f(y_1|\mu) \cdots f(y_n|\mu) \\
&= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y_i-\mu)^2}{2\sigma^2}\right] \\
&= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right]
\end{aligned}$$

故由 Neyman-Pearson 引理, MPT 的

$$\mathrm{RR} = \left\{ \frac{L(\mu_0)}{L(\mu_a)} < k \right\}$$

其中 k 爲由 α 待定的常數. 將 μ_0 與 μ_a 代入 $L(\mu)$, 得

$$\frac{L(\mu_0)}{L(\mu_a)} < k$$

等價於

$$\frac{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu_0)^2\right]}{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu_a)^2\right]} < k$$

整理化簡, 得

$$\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (y_i-\mu_0)^2 - \sum_{i=1}^n (y_i-\mu_a)^2\right]\right\} < k$$

取 ln 並整理, 得

$$\sum_{i=1}^n (y_i-\mu_0)^2 - \sum_{i=1}^n (y_i-\mu_a)^2 > -2\sigma^2 \ln k$$

展開化簡, 得

$$\sum_{i=1}^{n} y_i^2 - 2\mu_0 n\overline{y} + n\mu_0^2 - \sum_{i=1}^{n} y_i^2 + 2\mu_a n\overline{y} - n\mu_a^2 > -2\sigma^2 \ln k$$

整理, 得

$$\overline{y}(\mu_a - \mu_0) > \frac{1}{2n}(-2\sigma^2 \ln k - n\mu_0^2 + n\mu_a^2)$$

因爲 $\mu_a > \mu_0$, 上式亦等價於

$$\begin{aligned} \overline{y} &> \frac{1}{2n(\mu_a - \mu_0)}(-2\sigma^2 \ln k - n\mu_0^2 + n\mu_a^2) \\ &\stackrel{\text{def}}{=} k'(\mu_a) \end{aligned} \tag{1}$$

某個因 μ_a 而定的常數, 因爲 σ^2, n, μ_0 與 μ_a 均爲已知常數.

接著, 由 α 決定明確的 $k'(\mu_a)$. 由 α 的定義, 得

$$\begin{aligned} \alpha &= P(\text{RR, 當 } H_0 \text{ 眞時}) \\ &= P(\overline{Y} > k'(\mu_a), \text{ 當 } \mu = \mu_0) \\ &= P\left(\frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}} > \frac{k'(\mu_a) - \mu_0}{\sigma/\sqrt{n}}, \text{ 當 } \mu = \mu_0\right) \\ &= P\left(N(0,1) > \frac{k'(\mu_a) - \mu_0}{\sigma/\sqrt{n}}\right) \end{aligned} \tag{2}$$

此乃因爲當 $\mu = \mu_0$ 時, $\frac{\overline{Y}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ 所致.

註. 由 (1) 式知, 對所有的 $\mu_a > \mu_0$, RR 的型式均相同, 不因 μ_a 而定.

又

$$P(N(0,1) > z_\alpha) = \alpha \tag{3}$$

故比較 (2) 式與 (3) 式, 可設定

$$\frac{k'(\mu_a) - \mu_0}{\sigma/\sqrt{n}} = z_\alpha$$

即

$$k'(\mu_a) = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \tag{4}$$

不因 μ_a 而定.

因此,

$$\text{RR} = \left\{ \overline{y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \right\}$$

給出檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu = \mu_a$$

某固定的 $\mu_a > \mu_0$ 的 MPT.

最後, 由 (4) 式, 得對所有的 $\mu_a > \mu_0$, 只有單一的 RR, 故

$$\text{RR} = \left\{ \overline{y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \right\}$$

等價於

$$\text{RR} = \left\{ z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \right\}$$

一 Z-檢定, 給出檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu_a > \mu_0$$

的 UMPT, 如圖示.

註 1. 類推, 在意義水準 α 下, 檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu_a < \mu_0$$

時的 UMPT 的拒絕域

$$\text{RR} = \left\{ \overline{y} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} \right\}$$

如圖示.

註 2. 檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu \neq \mu_0$$

雙尾檢定時, 無 UMPT, 因爲由例 2 與註 1, 得二個不同的 RR, 而不是單一的 RR. 因此, 雙尾 Z-檢定不是檢定

$$H_0 : \mu = \mu_0 \text{ 對立於 } H_a : \mu \neq \mu_0$$

的 UMPT.

註 3. 問. 如何根據由僅含一未定參數 θ 值的母體所得的樣本, 求

$$H_0 : \theta \leq \theta_0 \text{ (複合) 對立於 } H_a : \theta > \theta_0 \text{ (複合)}$$

的檢定?

答. 假設已求得檢定

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

的 UMPT 所對應的 RR.

對於 $\theta_1 < \theta_0$, 若採用此 RR (即上述的 UMPT) 檢定

$$H_0'' : \theta = \theta_1 \text{ 對立於 } H_a : \theta > \theta_0$$

則一般而言,

$$\begin{aligned} \alpha_1 &\stackrel{\text{def}}{=} P(\text{RR, 當 } \theta = \theta_1) \\ &\leq \alpha = P(\text{RR, 當 } \theta = \theta_0) \end{aligned}$$

且對所有的 $\theta_a > \theta_0$,

$$\beta_1(\theta_a) = P(\text{RR}^c,\ 當\ \theta = \theta_a) = \beta(\theta_a)$$

即 power(θ_a) 不變, 如圖示. 也就是說, 以同樣的檢定 (即 RR) 對 H_0'' 及 H_a 檢定時, 甚至會更好. 因此, 定義檢定

$$H_0 : \theta \leq \theta_0\ 對立於\ H_a : \theta > \theta_0$$

的意義水準 α 爲 power(θ_0) (即 $H_0' : \theta = \theta_0$ 的 α) 且 UMPT 爲 $H_0' : \theta = \theta_0$ 的 RR.

註 4. 若母體除了欲檢定的參數 θ 外, 亦含有其它未確定的參數, 稱作干擾 (nusiance) 參數, 則不可使用 Neyman-Pearson 引理檢定 θ, 如

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

其中 σ^2 未知, 即干擾參數, 則 $H_0 : \mu = \mu_0$ 無法唯一地確定分布, 即不爲簡單的, 因而不可使用 Neyman-Pearson 引理. 所幸, 在 §10.11 會探討一個更一般且被廣泛採用, 特別是含干擾參數時更爲有用的檢定法, 稱作概似比值檢定 (likelihood ratio test).