單元 19: Rao-Blackwell 定理與最小變異不偏估計量
(課本 §9.5)

問. 如何找較小變異的不偏估計量?

答. 可由充分統計量的函數著手, 一個理論基礎爲

Rao-Blackwell 定理. 令 $\widehat{\theta}$ 是 θ 的不偏估計量且

$$\text{Var}(\widehat{\theta}) < \infty$$

若 U 是 θ 的充分統計量且定義

$$\widehat{\theta}^* = E(\widehat{\theta}|U)$$

乃一充分統計量 U 的函數, 則對所有的 θ,

$$E(\widehat{\theta}^*) = \theta$$

即 $\widehat{\theta}^*$ 是不偏的, 且

$$\text{Var}(\widehat{\theta}^*) \le \text{Var}(\widehat{\theta})$$

即 $\widehat{\theta}^*$ 比 $\widehat{\theta}$ 有較小的變異.

<證> 因爲 U 是 θ 的充分統計量, 根據定義, 給定 U, 任何統計量的條件分布 (包含 $\widehat{\theta}$) 均不因 θ 而定, 故對所

有的 θ, $\widehat{\theta}^* = E(\widehat{\theta}|U)$ 不會是 θ 的函數, 而可成爲一個 θ 的統計量.

接著, 根據條件期望值的性質 (定理 5.14) 及 $\widehat{\theta}$ 的不偏性, 得

$$E(\widehat{\theta}^*) = E[E(\widehat{\theta}|U)] = E(\widehat{\theta}) = \theta$$

故 $\widehat{\theta}^*$ 是 θ 的不偏估計量.

再由條件變異數的公式 (定理 5.15), 得

$$\begin{aligned} \text{Var}(\widehat{\theta}) &= E[\text{Var}(\widehat{\theta}|U)] + \text{Var}[E(\widehat{\theta}|U)] \\ &= \text{Var}(\widehat{\theta}^*) + E[\text{Var}(\widehat{\theta}|\text{U})] \qquad (1) \end{aligned}$$

最後, 因爲對所有的 u,

$$\text{Var}(\widehat{\theta}|U = u) \geq 0$$

故得

$$E[\text{Var}(\widehat{\theta}|U)] \geq 0$$

並由 (1) 式知,

$$\text{Var}(\widehat{\theta}^*) \leq \text{Var}(\widehat{\theta})$$

得證.

註 1 Rao-Blackwell 定理暗示有較小變異的不偏估計量會是 (或可被形成) 一充分統計量的函數.

註 2. 若 $\hat{\theta}$ 是不偏的, 改良的方式爲, 將 Rao-Blackwell 定理條件化在一充分統計量 U 上, 得

$$\hat{\theta}^* = E(\hat{\theta}|U) \overset{\text{記爲}}{=} h(U)$$

乃一 U 的函數, 有較小的變異數的不偏估計量.

問. 若再應用此定理以 U 條件化 $h(U)$ 上, 會得到更好的估計量嗎?

答. 因爲一般而言,

$$E[h(U)|U] = h(U) = \hat{\theta}$$

還是第一次應用後的結果, 無改進; 除非連續地條件化在不同的充分統計量上. 因此, 若一開始就選對了條件化的充分統計量, 就不需要連續地應用 Rao-Blackwell 定理條件化在不同的充分統計量上.

註 3. 此種對的充分統計量稱作最小充分統計量 (minimal sufficient statistic), 乃是一種可以盡可能地摘要或縮減樣本的充分統計量.

問. 如何求最小充分統計量?

答. 可根據下述的方法求得.

Lehmann 與 Scheffé 法. 令 $Y_1, \ldots, Y_n$ 爲一隨機樣本且未知參數爲 θ.

(i) 設定概似比值 (ratio of likelihoods)

$$\frac{L(x_1, \ldots, x_n|\theta)}{L(y_1, \ldots, y_n|\theta)}$$

(ii) 求一函數 $g(x_1, \ldots, x_n)$ 使得概似比值

$$\frac{L(x_1, \ldots, x_n|\theta)}{L(y_1, \ldots, y_n|\theta)}$$

不含 θ (free of θ) 若且唯若

$$g(x_1, \ldots, x_n) = g(y_1, \ldots, y_n)$$

(iii) $g(Y_1, \ldots, Y_n)$ 就是 θ 的一個最小充分統計量.

註 4. 針對本書所探討的分布, 由分解準則所確認的充分統計量 U, 一般而言, 就是最小充分統計量. 此外, 這些

最小充分統計量亦具有完備性 (completeness), 使得應用 Rao-Blackwell 定理時, 可保證

$$\widehat{\theta}^* = E(\widehat{\theta}|U)$$

在所有 θ 的不偏統計量中, 有最小的變異數, 即 $\widehat{\theta}^*$ 爲一 MVUE (minimum-variance unbiased estimator).

註 5. 直接求條件期望值 $E(\widehat{\theta}|U)$ 通常不易, 所幸針對所探討的分布, 若最小充分統計量 U 的某一函數

$$h(U)$$

是不偏的, 則 $h(U)$ 就是 MVUE.

例 1. 令 $Y_1, \ldots, Y_n$ 是一成功機率爲 p 的 Bernoulli 試驗. 試求一 p 的最小充分統計量以及 p 的一 MVUE.

<證> 由註 5 知, 先以 Lehmann 與 Scheffé 法 (或分解準則, 參看課本) 求最小充分統計量, 再求此統計量的某一函數使其爲 p 的不偏估計量, 即得一 MVUE.

首先, 由

$$P(Y_i = y_i) = p^{y_i}(1-p)^{1-y_i},\ y_i = 0, 1$$

以及獨立性, 得概似比值

$$\frac{L(x_1,\ldots,x_n|p)}{L(y_1,\ldots,y_n|p)} = \frac{P(Y_1=x_1,\ldots,Y_n=x_n|p)}{P(Y_1=y_1,\ldots,Y_n=y_n|p)}$$

$$= \frac{p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}}{p^{\sum_{i=1}^n y_i}(1-p)^{n-\sum_{i=1}^n y_i}} = \frac{\left(\frac{p}{1-p}\right)^{\sum_{i=1}^n x_i}}{\left(\frac{p}{1-p}\right)^{\sum_{i=1}^n y_i}}$$

$$= \left(\frac{p}{1-p}\right)^{\sum_{i=1}^n x_i - \sum_{i=1}^n y_i}$$

接著, 爲使 $L(x_1,\ldots,x_n|p)/L(y_1,\ldots,y_n|p)$ 不含 p 可設定

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

意即, 可取

$$g(x_1,\ldots,x_n) = \sum_{i=1}^n x_i$$

而得

$$g(x_1,\ldots,x_n) = g(y_1,\ldots,y_n)$$

因此, 由 Lehmann 與 Scheffé 法, 得

$$g(Y_1,\ldots,Y_n) = \sum_{i=1}^n Y_i$$

是 p 的最小充分統計量.

最後, 因爲

$$E\left(\sum_{i=1}^{n} Y_i\right) = np$$

故

$$\overline{Y} \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} Y_i$$

是最小充分統計量 $\sum_{i=1}^{n} Y_i$ 的函數且

$$E(\overline{Y}) = \frac{1}{n}(np) = p$$

即 $\overline{Y}$ 是不偏的, 得 $\overline{Y}$ 是一 MVUE.

例 2. 設隨機變數

$$Y_1, \ldots, Y_n \stackrel{\text{iid}}{\sim} f(y|\theta) = \begin{cases} \left(\frac{2y}{\theta}\right) e^{-y^2/\theta}, & y > 0 \\ 0, & \text{其它} \end{cases}$$

即 Weibull 分布的隨機樣本. 試求一 θ 的 MVUE.

<解> 如同例 1, 先求一 θ 的最小充分統計量, 再求此充分統計量的某一函數, 使其爲 θ 的不偏估計量, 即得一 θ 的 MVUE.

首先，由獨立及同分布，得概似比值

$$\frac{L(x_1,\ldots,x_n|\theta)}{L(y_1,\ldots,y_n|\theta)}=\frac{f(x_1|\theta)\cdots f(x_n|\theta)}{f(y_1|\theta)\cdots f(y_n|\theta)}$$

$$=\frac{\prod_{i=1}^n\left(\frac{2x_i}{\theta}\right)e^{-x_i^2/\theta}}{\prod_{i=1}^n\left(\frac{2y_i}{\theta}\right)e^{-y_i^2/\theta}}=\frac{x_1\cdots x_n e^{-\frac{1}{\theta}(\sum_{i=1}^n x_i^2)}}{y_1\cdots y_n e^{-\frac{1}{\theta}(\sum_{i=1}^n y_i^2)}}$$

$$=\frac{x_1\cdots x_n}{y_1\cdots y_n}e^{-\frac{1}{\theta}(\sum_{i=1}^n x_i^2-\sum_{i=1}^n y_i^2)}$$

接著，爲使上式不含 θ，可設定

$$\sum_{i=1}^n x_i^2=\sum_{i=1}^n y_i^2$$

意即可取

$$g(x_1,\ldots,x_n)=\sum_{i=1}^n x_i^2$$

而得

$$g(x_1,\ldots,x_n)=g(y_1,\ldots,y_n)$$

因此，由 Lehmann 與 Scheffé 法 (或參考課本根據分解準則)，得

$$g(Y_1,\ldots,Y_n)=\sum_{i=1}^n Y_i^2$$

是一 θ 的最小充分統計量.

又根據 $u = y^2/\theta$, $du = (2y/\theta)dy$ 的變數變換以及 gamma 函數的定義

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du, \ \alpha > 0$$

得

$$\begin{aligned} E(Y_i^2) &= \int_0^\infty y^2 \left(\frac{2y}{\theta}\right) e^{-y^2/\theta} dy \\ &= \theta \int_0^\infty u e^{-u} du = \theta \Gamma(2) = \theta \end{aligned}$$

由此得,

$$E\left(\sum_{i=1}^n Y_i^2\right) = \sum_{i=1}^n E(Y_i^2) = n\theta$$

故

$$\frac{1}{n}\sum_{i=1}^n Y_i^2$$

是最小充分統計量 $\sum_{i=1}^n Y_i^2$ 的函數且

$$E\left(\frac{1}{n}\sum_{i=1}^n Y_i^2\right) = \frac{1}{n}(n\theta) = \theta$$

即不偏的, 得 $\frac{1}{n}\sum_{i=1}^n Y_i^2$ 是 θ 的一 MVUE.

例 3. 設隨機變數

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

且 μ 與 σ^2 均未知. 試求 μ 與 σ^2 的 MVUE.

<解> (i) Lehmamm 與 Scheffé 法 (或分解準則, 參看課本). 首先, 由獨立同分布, 得概似比值

$$\begin{aligned}
\frac{L(x_1,\ldots,x_n|\mu,\sigma^2)}{L(y_1,\ldots,y_n|\mu,\sigma^2)} &= \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_i-\mu)^2/2\sigma^2}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-(y_i-\mu)^2/2\sigma^2}} \\
&= \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i-\mu)^2}} \\
&= \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^n (x_i-\mu)^2 - \sum_{i=1}^n (y_i-\mu)^2\right]\right\} \\
&= \exp\left\{-\frac{1}{2\sigma^2}\left[\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right)\right.\right. \\
&\qquad \left.\left. -2\mu\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\right]\right\}
\end{aligned}$$

接著, 爲使上式不含 μ 與 σ^2, 可設定

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad \text{且} \quad \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$$

意即選取

$$g_1(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i$$

且

$$g_2(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^2$$

而得

$$g_1(x_1, \ldots, x_n) = g_1(y_1, \ldots, y_n)$$

以及

$$g_2(x_1, \ldots, x_n) = g_2(y_1, \ldots, y_n)$$

因此,

$$\sum_{i=1}^{n} Y_i \quad \text{與} \quad \sum_{i=1}^{n} Y_i^2$$

聯合地形成

$$\mu \quad \text{與} \quad \sigma^2$$

的最小充分統計量.

(ii) 選取

$$\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

明顯地, $\sum_{i=1}^n Y_i$ 與 $\sum_{i=1}^n Y_i^2$ 的函數, 且

$$\begin{aligned} S^2 &= \frac{1}{n-1}\sum_{i=1}^n (Y_i-\overline{Y})^2 \\ &= \frac{1}{n-1}\left(\sum_{i=1}^n Y_i^2 - n\overline{Y}^2\right) \end{aligned}$$

亦爲 $\sum_{i=1}^n Y_i$ 與 $\sum_{i=1}^n Y_i^2$ 的函數.

又

$$E(\overline{Y}) = \mu \quad \text{且} \quad E(S^2) = \sigma^2$$

即分別爲 μ 與 σ^2 的不偏估計量.

因此, 樣本期望值 $\overline{Y}$ 爲 μ 的一 MVUE 且樣本變異數 S^2 爲 σ^2 的一 MVUE.

例 4. 設隨機變數

$$Y_1,\ldots,Y_n \overset{\text{iid}}{\sim} f(y) = \begin{cases} \frac{1}{\theta}e^{-y/\theta}, & y>0 \\ 0, & \text{其它} \end{cases}$$

即指數分布的隨機樣本. 試求一 Var(Y_i) 的 MVUE.

<解> 因爲 Var$(Y_i)=\theta^2$, 故由 θ 的最小充分統計量著手並求此最小充分統計量的某一函數, 使其爲 θ^2 的不偏估計量即可.

(i) θ 的最小充分統計量. 首先, 概似比值

$$\begin{aligned}\frac{L(x_1,\ldots,x_n|\theta)}{L(y_1,\ldots,y_n|\theta)} &= \frac{\prod_{i=1}^n \frac{1}{\theta}e^{-x_i/\theta}}{\prod_{i=1}^n \frac{1}{\theta}e^{-y_i/\theta}} \\ &= \exp\left[-\frac{1}{\theta}\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\right]\end{aligned}$$

接著, 爲使上式不含 θ, 可設定

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

因此,

$$\sum_{i=1}^n Y_i$$

是一 θ 的最小充分統計量.

(ii) 選取

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$$

一 $\sum_{i=1}^n Y_i$ 的函數, 且

$$E(\overline{Y}) = E(Y_1) = \theta$$

即 θ 的不偏估計量.

因此, $\overline{Y}$ 是 θ 的一 MVUE.

接著, 需求一 $\sum_{i=1}^{n} Y_i$ 的函數, 使其爲 θ^2 的不偏估計量. 因爲 $\overline{Y}$ 是 θ 的不偏估計量, 故可考慮 $\overline{Y}^2$, 得

$$
\begin{aligned}
E(\overline{Y}^2) &= \mathrm{Var}(\overline{Y}) + [E(\overline{Y})]^2 = \frac{\mathrm{Var}(Y_i)}{n} + \theta^2 \\
&= \frac{\theta^2}{n} + \theta^2 = \left(\frac{n+1}{n}\right)\theta^2
\end{aligned}
$$

故

$$\left(\frac{n}{n+1}\right)\overline{Y}^2$$

是 θ^2 的不偏估計量且亦爲 $\sum_{i=1}^{n} Y_i$ 的函數.

因此, $\left(\frac{n}{n+1}\right)\overline{Y}^2$ 是一 $\theta^2 = \mathrm{Var}(Y_i)$ 的 MVUE.

例 5. 設一由 Weibull 機率密度函數

$$f(y) = \begin{cases} \frac{2y}{\theta} e^{-y^2/\theta}, & y > 0 \\ 0, & \text{其它} \end{cases}$$

所取得的大小爲 10 的隨機樣本爲

$$0.637,\ 1.531,\ 0.733,\ 2.256,\ 2.364$$

以及

$$1.601,\ 0.152,\ 1.826,\ 1.868,\ 1.126$$

試求一 θ 的 95% 信賴區間.

<解> 由例 2 知, $\sum_{i=1}^{n} Y_i^2$ 是一 θ 的最小充分統計量, 故可採用 $\sum_{i=1}^{n} Y_i^2$ 求一 θ 的樞紐量. 令 $W_i = Y_i^2$, 則由變數變換 W_i 的 pdf

$$\begin{aligned} f_{W_i}(w) &= f_{Y_i}(y)\frac{dy}{dw}\Big|_{y^2=w} \\ &= \frac{2y}{\theta}e^{-y^2/\theta}\frac{dy}{dw}\Big|_{y=\sqrt{w},\, w>0} \\ &= \frac{2\sqrt{w}}{\theta}e^{-w/\theta}\frac{1}{2\sqrt{w}} = \frac{1}{\theta}e^{-w/\theta},\ w>0 \end{aligned}$$

即

$$W_i \sim \exp(\theta) = \text{gamma}(1,\theta)$$

因 θ 而定, 不符樞紐量的條件. 接著, 令 $T_i = \frac{2W_i}{\theta}$, 則 T_i 的 pdf

$$\begin{aligned} f_{T_i}(t) &= f_{W_i}(w)\frac{dw}{dt}\Big|_{w=\frac{\theta}{2}t,\, t>0} \\ &= \frac{1}{\theta}e^{-w/\theta}\frac{dw}{dt}\Big|_{w=\frac{\theta}{2}t,\, t>0} \\ &= \frac{1}{\theta}e^{-t/2}\frac{\theta}{2} = \frac{1}{2}e^{-t/2},\ t>0 \end{aligned}$$

即

$$T_i = \frac{2W_i}{\theta} = \frac{2Y_i^2}{\theta} \sim \text{gamma}(1,2) = \chi^2(2)$$

不因 θ 而定. 故由 Y_i 的獨立性及同分布, 得

$$T_1, \ldots, T_{10} \overset{\text{iid}}{\sim} \chi^2(2)$$

以及

$$\sum_{i=1}^{10} T_i = \frac{2}{\theta} \sum_{i=1}^{10} Y_i^2 \sim \chi^2(20)$$

分布不因 θ 而定, 故可選取

$$\frac{2}{\theta} \sum_{i=1}^{10} Y_i^2$$

作爲 θ 的樞紐量.

若 a 與 b 使得

$$P\left(a < \frac{2}{\theta} \sum_{i=1}^{10} Y_i^2 < b\right) = 0.95$$

則經由整理, 得 θ 的 95% 信賴區間爲

$$\left(\frac{2\sum_{i=1}^{10} Y_i^2}{b}, \frac{2\sum_{i=1}^{10} Y_i^2}{a}\right)$$

由表 6,

$$P(\chi^2(20) > 34.170) = 0.025$$

且

$$P(\chi^2(20) > 9.591) = 0.975$$

故

$$a = 9.591 \quad \text{且} \quad b = 34.170$$

或 R:

$$\mathtt{a} = \mathtt{qchisq}(0.025, \mathtt{df} = 20)$$

且

$$\mathtt{b} = \mathtt{qchisq}(0.975, \mathtt{df} = 20)$$

另

$$\sum_{i=1}^{10} Y_i^2 = 24.643$$

因此, θ 的 95% 信賴區間爲

$$\left(\frac{2(24.643)}{34.170}, \frac{2(24.643)}{9.591}\right) = (1.442, 5.139)$$