單元 18: 充分性
(課本 §9.4)

設 $Y_1, \ldots, Y_n$ 是一大小爲 n 的隨機樣本, 則 $\overline{Y}$ 是 μ 的不偏估計量以及 S^2 是 σ^2 的不偏估計量.

上述處理資料的程序乃將大小爲 n 的隨機樣本

$$Y_1, \ldots, Y_n \overset{\text{(摘要成, 縮減成)}}{\Rightarrow} \overline{Y} \text{ 與 } S^2$$

二值.

問. 是否會流失有關 μ 與 σ^2 的訊息?

定義 (非正式, 直觀). 一統計量 (statistic) 稱作充分的 (sufficient) 若此統計量可由樣本中摘要出所有與目標參數相關的訊息.

舉例 (充分統計量的概念, 想法), 設 $X_1, \ldots, X_n$ 是成功機率爲 p 的 Bernoulli 試驗. 則

$$Y = \sum_{i=1}^{n} X_i$$

即 n 次試驗中的成功數, 可給出一些有關 p 的訊息.

問. 若知道 Y 的值, 是否可由觀察其它的 $X_1, \ldots, X_n$ 的函數, 而獲得更多有關 p 的訊息?

答. 一個方式是計算給定 Y 時, $X_1, \ldots, X_n$ 的條件分布, 得

$$
\begin{aligned}
& P(X_1 = x_1, \ldots, X_n = x_n | Y = y) \\
= \;& \frac{P(X_1 = x_1, \ldots, X_n = x_n, Y = y)}{P(Y = y)} \\
= \;& \begin{cases} \dfrac{p^y(1-p)^{n-y}}{\binom{n}{y} p^y (1-p)^{n-y}} = \dbinom{n}{y}, & \text{若 } \displaystyle\sum_{i=1}^{n} x_i = y \\ 0, & \text{其它} \end{cases}
\end{aligned}
$$

不因 p 而定, 意即, 若知道 Y 的值, 無法經由觀察其它 $X_1, \ldots, X_n$ 的函數, 而得到有關 p 的額外訊息. 故 Y 包含了所有與 p 相關的訊息. 這說明了 $X_1, \ldots, X_n$ 所提供的成功總數的訊息 Y 以外, 其它的訊息, 如哪幾次 (何時) 是成功等等, 對成功機率 p 而言, 都是不相關的.

定義 (正式). 令 $Y_1, \ldots, Y_n$ 爲一由含未知參數 θ 的機率分布所取得的隨機樣本. 稱統計量

$$U = g(Y_1, \ldots, Y_n)$$

對 θ 是充分的 (sufficient) 若給定 U 時, $Y_1, \ldots, Y_n$ 的條件分布不因 θ 而定 (not depending on θ, 不取決於 θ).

定義. 令 $y_1, \ldots, y_n$ 爲分布函數因 θ 而定的隨機變數 $Y_1, \ldots, Y_n$ 的樣本觀察值. 若 $Y_1, \ldots, Y_n$ 爲離散隨機變數, 則此樣本的概似值 (likelihood of the sample)

$$L(y_1, \ldots, y_n|\theta) \overset{\text{def}}{=} P(Y_1 = y_1, \ldots, Y_n = y_n)$$

即 $y_1, \ldots, y_n$ 的聯合機率 (joint probability). 又若 $Y_1, \ldots, Y_n$ 爲連續隨機變數, 則此樣本的概似值

$$L(y_1, \ldots, y_n|\theta) \overset{\text{def}}{=} f(y_1, \ldots, y_n)$$

即 $y_1, \ldots, y_n$ 的聯合密度 (joint density). 均簡記成 $L(\theta)$.

定理 (分解準則, factorization criterion). 令 U 爲一隨機樣本 $Y_1, \ldots, Y_n$ 的統計量. 則 U 是 θ 的充分統計量若且唯若概似值

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

其中 $g(u, \theta)$ 僅爲 u 與 θ 的函數且 $h(y_1, \ldots, y_n)$ 不爲 θ 的函數.

<證> 超出本書範圍, 略.

例. 設

$$Y_1, \ldots, Y_n \sim \exp(\theta)$$

試證 $\overline{Y}$ 是 θ 的充分統計量.

<證> 首先, Y_i 的 pdf

$$f(y_i) = \begin{cases} \frac{1}{\theta} e^{-y_i/\theta}, & 0 \le y_i < \infty \\ 0, & \text{其它} \end{cases}$$

故, 根據 $Y_1, \ldots, Y_n$ 的相互獨立性及同分布, 此樣本的概似值

$$\begin{aligned} L(\theta) &= L(y_1, \ldots, y_n|\theta) = f(y_1, \ldots, y_n|\theta) \\ &= f(y_1|\theta) f(y_2|\theta) \cdots f(y_n|\theta) \\ &= \frac{e^{-y_1/\theta}}{\theta} \frac{e^{-y_2/\theta}}{\theta} \cdots \frac{e^{-y_n/\theta}}{\theta} \\ &= \frac{e^{-\sum_{i=1}^n y_i/\theta}}{\theta^n} = \frac{1}{\theta^n} e^{-n\overline{y}/\theta} \end{aligned}$$

乃一 θ 與 $\overline{y}$ 的函數.

因此, 取

$$g(\overline{y}, \theta) = \frac{1}{\theta^n} e^{-n\overline{y}/\theta}$$

以及

$$h(y_1, \ldots, y_n) = 1$$

得

$$L(y_1, \ldots, y_n|\theta) = g(\overline{y}, \theta) h(y_1, \ldots, y_n)$$

且由分解準則知, $\overline{Y}$ 是 θ 的充分統計量.