

## 第五章

# 資料減縮

### 5.2 充分統計量

在上一章我們介紹一些將資料做初步整理的圖示法。那些是屬於敘述統計的範疇，並非本書主要探討的。本節我們以數理統計的觀點來討論關於資料的處理。

對所收集到的資料，有沒有必要一直全部攜帶(或展示)著？有時對如此多的資料，也很難理解。是否可扼要地掌握資料的一些特性即可？通常我們是藉由計算**統計量**(statistic)來達到此目的。

設有一組隨機樣本 $X_1, \dots, X_n$ 。若令 $\mathbf{X} = (X_1, \dots, X_n)$ ，則 $\mathbf{X}$ 的一實值函數(也可以是向量值) $T(\mathbf{X}) = T(X_1, \dots, X_n)$ ，便稱為 $\mathbf{X}$ 之一**統計量**。統計量仍為一隨機變數(或隨機向量)，它的定義可說是非常寬鬆，唯一的限制是函數 $T$ 要與參數 $\theta$ 無關。諸如樣本平均

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i,$$

樣本變異數

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2,$$

## 2 第五章 資料減縮

最小的順序統計量 $X_{(1)}$ ，最大的順序統計量 $X_{(n)}$ ，中位數 $\text{med}X$ ，第一個樣本 $X_1$ ，常數5， $S(\mathbf{X}) = (X_{(1)}, X_{(n)})$ ，甚至 $T(\mathbf{X}) = (X_1, \dots, X_n)$ ，皆為統計量。但 $\overline{X}_n + \theta$ 則非統計量。在較一般的情況，即使 $X_1, \dots, X_n$ 不為i.i.d.，它們的一個函數，我們仍稱為統計量。

某些統計量可保留資料中有用的資訊，而去掉的都是不相干的，不影響對參數的推估。我們常說統計的工作，就是要收集資料、整理資料等。這種統計量可以達到**資料減縮**(data reduction)的目的，並使資訊毫無損失，是一種高層次的整理資料，自然是我們所喜愛的。我們稱之為**充分統計量**(sufficient statistic)。

在給充分統計量的定義前，我們先藉下例來說明此概念之內涵。

**例2.1** 我們有時會想估計一銅板出現正面的機率，或某種新藥的治癒率。對這類的問題，若以1表成功，0表失敗，每次成功的機率以 $\theta$ 表之，則經 $n$ 次試驗後，可得一組i.i.d.之 $\text{Ber}(\theta)$ 隨機變數 $X_1, \dots, X_n$ 。基於所觀察到的 $X_1 = x_1, \dots, X_n = x_n$ ，我們想對 $\theta$ 做一些推估。

直觀上我們會認為，如果知道 $x_1, \dots, x_n$ 中共有幾個1，也就是知道 $x_1, \dots, x_n$ 之和，則即使再獲知1在那幾次試驗發生，並無法對 $\theta$ 做出更精確的推估。如何驗證此直觀是否正確呢？令統計量

$$T = T(\mathbf{X}) = X_1 + \dots + X_n$$

表總共之成功數。對一固定的 $t$ ， $t = 0, 1, 2, \dots, n$ ，對應 $T = t$ ，共有 $\binom{n}{t}$ 種可能的 $x_1, \dots, x_n$ 。給定 $T = t$ ，求出此 $\binom{n}{t}$ 個會有 $t$ 個成功發生之不同事件的機率。若有些 $\theta$ ，使得某些 $t$ 個成功發生的事件之機率，與其他的 $t$ 個成功發生的事件之機率不同，則知道 $t$ 個成功在何處發生，比僅知道 $t$ 個成功，便提供較多的資訊。反之，若所有 $t$ 個成功發生的事件之機率都相同，則顯然 $t$ 個成功發生之位置便無關緊要，而總共之成功數 $t$ ，便對 $\theta$ 提供所有資訊。

首先有

$$P(X_1 = x_1, \dots, X_n = x_n | T = t)$$

$$\begin{aligned}
&= \frac{P(X_1 = x_1, \dots, X_n = x_n, T = t)}{P(T = t)} \\
&= \frac{P(X_1 = x_1, \dots, X_n = x_n)}{P(T = t)}, \quad \text{若 } \sum_{i=1}^n x_i = t,
\end{aligned}$$

否則為0。再利用 $T$ 有 $\mathcal{B}(n, \theta)$ 分佈, 上述條件機率又等於

$$\begin{aligned}
\frac{\prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} &= \frac{\theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} \\
&= \frac{1}{\binom{n}{t}}, \quad \text{若 } \sum_{i=1}^n x_i = t,
\end{aligned}$$

否則為0。即得證總共之成功數 $T$ , 對 $\theta$ 提供所有可能的資訊。

在上例中,  $\mathbf{X}$ 之p.d.f.

$$\theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i},$$

可分解成兩部分: 其一為給定 $T = t$ 之下,  $\mathbf{X}$ 之條件機率

$$\binom{n}{t}^{-1}, \quad \text{若 } \sum_{i=1}^n x_i = t,$$

其二為 $T = t$ 之機率

$$\binom{n}{t} \theta^t (1-\theta)^{n-t}, t = \sum_{i=1}^n x_i.$$

前者與 $\theta$ 無關, 所以未能提供對 $\theta$ 之任何資訊, 而後者包含資料中所有對 $\theta$ 之資訊。所以若要對 $\theta$ 做推估, 合理地說, 只要知道 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 就夠了。而此 $T(\mathbf{X})$ 當然比全部資訊 $\mathbf{X} = (X_1, \dots, X_n)$ 輕便多了。

**定義2.1** 設 $\mathbf{X}$ 之分佈與參數 $\theta$ 有關。對一統計量 $T(\mathbf{X})$ , 若給定 $T(\mathbf{X})$ ,  $\mathbf{X}$ 之條件分佈“與 $\theta$ 無關”(does not depend on  $\theta$ ), 則 $T(\mathbf{X})$ 稱為 $\theta$ 之一充分統計量。

若 $T$ 為 $\theta$ 之一充分統計量, 則知道 $T$ 之後, 便掌握對 $\theta$ 之所有資訊(即知道 $T$ 之後,  $\mathbf{X}$ 之變化已不再仰賴 $\theta$ ), 也就是樣本中已不再含有更多關於 $\theta$ 之資訊。知道樣本的其他資訊, 對 $\theta$ 的了解不再有幫助。在例2.1中, 知道 $\sum_{i=1}^n X_i = t$ , 則可決定任一 $X_1 = x_1, \dots, X_n = x_n$ 事件之機率(不再仰賴 $\theta$ ), 不須再多知道諸如 $X_1$ 及 $X_3$ 之值。可以這麼說, 剩餘的資訊中, 對了解 $\theta$ 已不具任何價值了。這是“充分”此一名稱的由來。投擲一銅板 $n$ 次, 我們只要知道共得幾個正面, 而不須知道各次的結果, 便已掌握所有推估出現正面機率之資訊, 資料明顯大幅度地減縮。

費雪在他於西元1922年的論文中, 首度提出充分統計量的概念。隨後費雪並認為, 就估計的目的而言, 充分統計量等價於原始的資料, 因此好的估計量只須基於充分統計量(good estimators should depend on the original observations only through sufficient statistics)。後來此原則被擴大, 大部分的統計學家都接受, 不論是估計、檢定(test)及信賴區間(confidence interval)等對參數的推論, 均只須基於充分統計量。這就是所謂**充分性原則**(Sufficient principle)。亦即若 $T(\mathbf{X})$ 為 $\theta$ 之一充分統計量, 則對任兩組觀察值 $\mathbf{x}$ 及 $\mathbf{y}$ , 只要 $T(\mathbf{x}) = T(\mathbf{y})$ , 則不論觀察到 $\mathbf{X} = \mathbf{x}$ 或 $\mathbf{X} = \mathbf{y}$ , 對 $\theta$ 之推論均須相同。不過必須一提的是, 還是難免有些“好的”的統計量並不基於充分統計量, 見Moore(1971)及Levy(1985)。

充分統計量一定存在。明顯可看出, 對一組隨機變數 $X_1, \dots, X_n$ , 且其分佈與參數 $\theta$ 有關, 統計量 $T = (X_1, \dots, X_n)$ , 必為 $\theta$ 之一充分統計量。此因

$$P(X_1 = x_1, \dots, X_n = x_n | T = (x_1, \dots, x_n)) = 1,$$

且對 $(a_1, \dots, a_n) \neq (x_1, \dots, x_n)$ ,

$$P(X_1 = a_1, \dots, X_n = a_n | T = (x_1, \dots, x_n)) = 0。$$

故給定 $T$ 之後,  $X_1, \dots, X_n$ 之條件分佈與 $\theta$ 無關。由於充分統計量是為減縮資料(但不減少對參數 $\theta$ 之資訊), 此組天然的充分統計量顯然並不好。並未真正減縮資料。但順序統計量 $(X_{(1)}, \dots, X_{(n)})$ 必為 $\theta$ 之一充分

統計量(習題第5題)。而 $(X_{(1)}, \dots, X_{(n)})$ 當然較 $(X_1, \dots, X_n)$ 更能減縮資料。此因有 $n!$ 組 $(X_1, \dots, X_n)$ 皆對應同一組 $(X_{(1)}, \dots, X_{(n)})$ 。所以的確存在真正能減縮資料的充分統計量。由此看來似乎充分統計量還可以比好壞：最能減縮資料的充分統計量應視為“最佳的”。我們給下述定義。

**定義2.2** 設有一充分統計量 $T(\mathbf{X})$ ，若對每一充分統計量 $T_1(\mathbf{X})$ ， $T(\mathbf{X})$ 皆為 $T_1(\mathbf{X})$ 之函數，則稱 $T(\mathbf{X})$ 為**最小充分統計量**(minimal sufficient statistic)。

當 $T$ 為一最小充分統計量，表示資料無法再縮減，而不會失去充分性。底下定理可用來判別一函數是否為另一函數之函數。只要由函數之定義立即可得知其證法，因此我們略去證明。

**定理2.1** 設 $f, g$ 為二有相同定義域 $D$ 之函數。則 $f$ 為 $g$ 之函數，若且唯若下述條件成立：若 $g(x_1) = g(x_2)$ ，則 $f(x_1) = f(x_2)$ ， $x_1, x_2 \in D$ 。

由定理2.1知，若函數 $f$ 與 $g$ 有相同定義域，且 $g$ 為嚴格單調，則 $f$ 為 $g$ 之函數(習題第1題)。

**例2.2** 設 $f(x) = x^3$ ， $g(x) = x^2$ ， $x \in R$ ，則因 $g(-2) = g(2)$ ，而 $f(2) \neq f(-2)$ ，故 $f$ 非 $g$ 之函數。但顯然 $g$ 為 $f$ 之函數： $g(x) = \sqrt[3]{(f(x))^2}$ 。不過如果定義域改為 $[0, \infty)$ ，則 $f$ 便為 $g$ 之函數了： $f(x) = \sqrt{(g(x))^3}$ 。此時 $g$ 當然還是 $f$ 的函數： $g(x) = \sqrt[3]{(f(x))^2}$ 仍成立。

不過有時並非那麼容易看出 $f$ 是否為 $g$ 之函數。如 $f(x) = x^3 + 3x^2 + 1$ ， $g(x) = x^4 - 2x^2 + 1$ ， $x \geq 0$ 。事實上此時 $f$ 並非 $g$ 之函數(習題第3題)。

一統計量為樣本之一函數(可能是多維)，既然是函數就有資料減縮的功能。除非是樣本的1-1函數，沒有達到資料減縮的目的。譬如說有樣本 $\mathbf{X} = (X_1, \dots, X_n)$ ，則諸如 $T_1(\mathbf{X}) = (X_1 + 1, \dots, X_n + 1)$ ， $T_2(\mathbf{X}) = (X_1^3, \dots, X_n^3)$ ， $T_3(\mathbf{X}) = (e^{X_1}, \dots, e^{X_n})$ ，皆為與樣本1-1對應

之統計量。此二統計量所攜帶資料的“厚重”程度，與樣本 $\mathbf{X}$ 當然是一樣的。除此之外，不為樣本之 $1-1$ 函數的統計量，皆會減縮資料。譬如說 $T_4(\mathbf{X}) = (X_1^2, \dots, X_n^2)$ ，且 $X_i$ 取值在 $R$ 上，則 $T_4$ 將資料減縮了： $(1, \dots, 1)$ 與 $(-1, \dots, -1)$ 對應到相同的 $(1, \dots, 1)$ 。

上述討論指出，若有 $U, T$ 二統計量，其中 $T$ 為一充分統計量，且 $T = H(U)$ ，即 $T$ 為 $U$ 之一函數，則 $U$ 亦為一充分統計量：給定 $U$ 便給定 $T$ ，而給定 $T$ ， $\mathbf{X}$ 之條件分佈與參數 $\theta$ 無關。尤有進者，除非 $H$ 是 $U$ 的 $1-1$ 函數，否則 $T$ 較 $U$ 更能減縮資料。

二統計量 $T_1, T_2$ ，若 $T_2$ 為 $T_1$ 之 $1-1$ 函數，便稱 $T_1$ 與 $T_2$ 等價。此時知道 $T_1$ 便可唯一決定 $T_2$ ，知道 $T_2$ 也可唯一決定 $T_1$ 。二等價統計量本質上是沒有差別的。也可以定理2.1的想法來判定等價：二統計量 $T_1, T_2$ ，若滿足“ $T_1(\mathbf{x}) = T_1(\mathbf{y})$ ，若且唯若 $T_2(\mathbf{x}) = T_2(\mathbf{y})$ ”，便為等價。等價的統計量對參數提供同樣多的資訊，且達到相同程度的資料減縮效果。因此二等價統計量必同時為充分統計量或同時不是，也必同時為最小充分統計量或同時不是。由此可知最小充分統計量若存在，便不唯一。又等價統計量也是一種等價關係，也就是有反身性、對稱性及遞移性。即設有三個統計量 $T_1, T_2$ 及 $T_3$ 。則 $T_1$ 與 $T_1$ 等價；若 $T_1$ 與 $T_2$ 等價，則 $T_2$ 與 $T_1$ 等價；若 $T_1$ 與 $T_2$ 等價，且 $T_2$ 與 $T_3$ 等價，則 $T_1$ 與 $T_3$ 等價。

一般而言，資料減縮有兩方面的意義，一個是維數減少，如由 $(X_1, \dots, X_n)$ 至 $\sum_{i=1}^n X_i$ ，維數由 $n$ 維減至1維；一個是經由非 $1-1$ 的變換(或說函數對應)，如取值在實數，且由 $(X_1, \dots, X_n)$ 至 $(X_1^2, \dots, X_n^2)$ 。當然第一種意義也屬於非 $1-1$ 的變換。簡言之，二等價統計量，減縮資料的效果是一樣的(由此也可知最小充分統計量不唯一)；而一統計量之非 $1-1$ 的函數，有最佳的減縮資料的效果。最小充分統計量，由於要達到最大資料減縮的效果，因此它是任一充分統計量的函數。

**例2.3** 設有統計量 $T_1 = \sum_{i=1}^n X_i$ ， $T_2 = \sum_{i=1}^n X_i^2$ ， $T_3 = \sum_{i=1}^n X_i^3$ 。是否有那二統計量等價？

**解.**  $T_1$ 與 $T_2$ 一般而言是不等價的。以 $n=2$ 為例：

$$1^2 + (-1)^2 = 1^2 + 1^2 = 2, \text{ 但 } 1 + (-1) \neq 1 + 1。$$

即使 $X_i$ 皆取正值,  $T_1$ 與 $T_2$ 也不等價:

$$7^2 + 6^2 = 9^2 + 2^2 = 85, \text{ 但 } 7 + 6 \neq 9 + 2。$$

$T_1$ 與 $T_3$ 也不等價, 仍以 $n = 2$ 為例:

$$12^3 + 1^3 = 10^3 + 9^3 = 1729, \text{ 但 } 12 + 1 \neq 10 + 9。$$

$T_2$ 與 $T_3$ 也不等價,  $n$ 仍取為2:

$$12^3 + 1^3 = 10^3 + 9^3, \text{ 但 } 12^2 + 1^2 \neq 10^2 + 1^2。$$

爲了減縮資料, 我們有各種統計量, 但又不能過度減縮, 否則留下來的資訊便不夠多了。例如, 若 $T(\mathbf{X}) = 3$ , 當然是非常精簡了, 但此統計量可說是毫無用途。爲了參數推估的目的, 我們將統計量減至充分統計量的地步, 而在所有充分統計量中, 我們又偏好其中最精減者, 也就是**最小充分統計量**。由定義2.2知, 最小充分統計量爲任一充分統計量的函數, 因此給定某一充分統計量, 最小充分統計量若不是與此充分統計量等價, 就是攜帶的資料的確較其減縮。

在例2.1中,  $T = \sum_{i=1}^n X_i$ ,  $S = T/n = \bar{X}_n$ ,  $U = T/3$ ,  $V = 2T - 3$ 等, 皆爲參數 $\theta$ 之充分統計量。事實上這幾個統計量爲等價。二等價的統計量必同時爲充分統計量或同時非充分統計量。又若 $X$ 有 $\mathcal{N}(0, \sigma^2)$ 分佈, 且設樣本數 $n = 1$ , 則顯然 $T_1(X) = X$ ,  $T_2(X) = X^2$ ,  $T_3(X) = |X|$ 等, 皆爲 $\sigma^2$ 之充分統計量, 其中 $T_2$ 與 $T_3$ 等價。 $T_2$ 與 $T_3$ 皆爲 $\sigma^2$ 之最小充分統計量。

**例2.4** 設 $X, Y$ 爲i.i.d., 皆有 $\mathcal{E}(\theta)$ 分佈,  $\theta > 0$ 。令 $Z = X + Y$ , 則 $Z$ 有 $\Gamma(2, 1/\theta)$ 分佈, 即 $Z$ 之p.d.f. $f_Z(z)$ 爲

$$f_Z(z) = \theta^2 z e^{-\theta z}, z > 0。$$

利用變數代換可得 $X, Z$ 之聯合p.d.f.爲

$$f_{X,Z}(x, z) = \theta^2 e^{-\theta z}, 0 < x < z。$$

故給定  $X + Y = z$ ,  $X$  之條件 p.d.f. 為

$$f_{X|Z}(x|z) = \frac{f_{X,Z}(x, z)}{f_Z(z)} = \frac{1}{z}, 0 < x < z.$$

即得在給定  $X + Y = z$  之下,  $(X, Y) \stackrel{d}{=} (U, V)$ , 其中  $U$  有  $\mathcal{U}(0, z)$  分佈,  $V = z - U$ 。故依定義知,  $X + Y$  為  $\theta$  之一充分統計量。

**例2.5** 設隨機變數  $X$  取值在  $\{1, 2, 3, 4\}$ , 其 p.d.f. 為

$$f(x) = \begin{cases} 1/4 & , x = 1, 2, \\ (1 + \theta)/4 & , x = 3, \\ (1 - \theta)/4 & , x = 4, \end{cases}$$

其中  $\theta \in [0, 1]$ 。若令  $T(1) = T(2) = 1, T(3) = 3, T(4) = 4$ , 則  $P(T = 1) = 1/2, P(T = 3) = (1 + \theta)/4, P(T = 4) = (1 - \theta)/4$ 。又

$$\begin{aligned} P(X = x|T = 1) &= \begin{cases} (1/4)/(1/2) = 1/2 & , x = 1, 2, \\ 0 & , x = 3, 4, \end{cases} \\ P(X = x|T = 3) &= \begin{cases} 1 & , x = 3, \\ 0 & , \text{其他}, \end{cases} \\ P(X = x|T = 4) &= \begin{cases} 1 & , x = 4, \\ 0 & , \text{其他}, \end{cases} \end{aligned}$$

因此  $T$  為  $\theta$  之一充分統計量。

讀者可否看出  $T$  為一最小充分統計量? 而可否給出所有最小充分統計量(習題第17題)?

**例2.6** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{N}(\theta, 1)$  分佈所產生之隨機樣本。我們可能會猜測  $\bar{X}_n = \sum_{i=1}^n X_i/n$  為  $\theta$  之一充分統計量。底下來看如何證明。

首先不難看出只要能證明  $T = \sum_{i=1}^n X_i$  為  $\theta$  之一充分統計量即可。考慮由  $X_1, \dots, X_n$  至  $T, Y_2, \dots, Y_n$  之一變數代換, 其中

$$Y_2 = X_2 - X_1, Y_3 = X_3 - X_1, \dots, Y_n = X_n - X_1.$$

由於

$$\begin{aligned} X_1 &= (T - Y_2 - \cdots - Y_n)/n, \\ X_2 &= Y_2 + X_1 = Y_2 + (T - Y_2 - \cdots - Y_n)/n, \\ &\vdots \\ X_n &= Y_n + X_1 = Y_n + (T - Y_2 - \cdots - Y_n)/n, \end{aligned}$$

故由  $X_1, \dots, X_n$  至  $T, Y_2, \dots, Y_n$  為 1-1 變換。若能證出在給定  $T = t$  之下,  $Y_2, \dots, Y_n$  之條件分佈與  $\theta$  無關, 則在給定  $T = t$  之下,  $T, Y_2, \dots, Y_n$  之條件分佈也與  $\theta$  無關。因此在給定  $T = t$  之下,  $X_1, \dots, X_n$  之條件分佈與  $\theta$  無關。由是便得證  $T$  為  $\theta$  之一充分統計量。

事實上,  $(T, Y_2, \dots, Y_n)$  有多變量常態分佈, 且  $Y_i$  與  $T$  為無相關,  $\forall i = 2, \dots, n$ : 由對稱性

$$\text{Cov}(T, Y_i) = \text{Cov}(T, X_i - X_1) = \text{Cov}(T, X_i) - \text{Cov}(T, X_1) = 0。$$

故  $(Y_2, \dots, Y_n)$  與  $T$  獨立。即得在給定  $T = t$  之下,  $Y_2, \dots, Y_n$  之條件分佈, 與非條件下  $(Y_2, \dots, Y_n)$  之分佈相同。而後者有  $n-1$  維常態分佈, 期望向量為  $(0, \dots, 0)$ :

$$E(Y_i) = E(X_i) - E(X_1) = \theta - \theta = 0,$$

共變異數矩陣為

$$\begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix}。$$

此因

$$\begin{aligned} \text{Cov}(Y_i, Y_i) &= \text{Var}(Y_i) = \text{Var}(X_i) + \text{Var}(X_1) = 2, \\ \text{Cov}(Y_i, Y_j) &= \text{Cov}(X_i - X_1, X_j - X_1) = \text{Cov}(X_i, X_j) \\ &= \text{Var}(X_1) = 1, \quad i \neq j。 \end{aligned}$$

即得證  $T$  為  $\theta$  之一充分統計量。

由上述幾個例子知, 要由充分統計量的定義, 來判定一統計量是否為充分, 並非很方便, 不但要先猜出那一個統計量為充分, 且必須經過一繁瑣的過程, 以檢驗猜測是否正確。此二步驟有時皆非容易。Halmos 及 Savage 在西元1949年, 給出一由p.d.f.來找出充分統計量的簡單方法, 稱為**分解定理**(Factorization theorem)。我們敘述此結果如下, 證明可參考 Casella and Berger(1990)一書pp.250-251。

**定理2.2** 令 $f(\mathbf{x}|\theta)$ 表 $\mathbf{X}$ 之聯合p.d.f.。則 $T(\mathbf{X})$ 為 $\theta$ 之一充分統計量, 若且唯若存在函數 $g(t|\theta)$ 及 $h(\mathbf{x})$ , 使得對所有樣本點 $\mathbf{x}$ 及參數 $\theta$ ,

$$(2.1) \quad f(\mathbf{x}|\theta) = g(T(\mathbf{x})|\theta)h(\mathbf{x})。$$

在上述定理中,  $g(t|\theta)$ 不一定為 $T(\mathbf{X})$ 之p.d.f., 但與 $T(\mathbf{X})$ 之p.d.f.的商, 為一 $\mathbf{x}$ 之函數。換句話說, 若 $T(\mathbf{X})$ 為 $\theta$ 之一充分統計量, 且 $T(\mathbf{X})$ 之p.d.f.存在, 則(2.1)式中之 $g$ 可取為 $T(\mathbf{X})$ 之p.d.f.。如前 $\theta$ 可能是一向量,  $T(\mathbf{X})$ 也可能為多維隨機向量。

**例2.7** 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{B}(m, p)$ 分佈所產生之隨機樣本,  $m$ 設為已知, 即 $\theta = p$ 。首先

$$f(\mathbf{x}|p) = p^{\sum_{i=1}^n x_i} (1-p)^{mn - \sum_{i=1}^n x_i} \prod_{i=1}^n \binom{m}{x_i}, x_i = 0, 1, \dots, m。$$

令 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ , 且取

$$g(t|p) = p^t (1-p)^{mn-t}, t = 0, 1, \dots, mn,$$

$$h(\mathbf{x}) = \prod_{i=1}^n \binom{m}{x_i}, x_i = 0, 1, \dots, m。$$

則由分解定理,  $T$ 為 $p$ 之一充分統計量。

**例2.8** 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本,  $\theta > 0$ 。則

$$f(\mathbf{x}|\theta) = \theta^n e^{-\theta \sum_{i=1}^n x_i}, x_i > 0, i = 1, \dots, n。$$

令  $T = \sum_{i=1}^n X_i$ , 且取  $g(t|\theta) = \theta^n e^{-\theta t}$ ,  $t > 0, \theta > 0$ ,  $h(\mathbf{x}) = 1$ ,  $x_i > 0, i = 1, \dots, n$ 。則由分解定理,  $T$  為  $\theta$  之一充分統計量。

與例2.4比較, 上例顯示分解定理對判定充分統計量, 提供很大的便利。如果做熟練, 並不需將  $g$  與  $h$  寫出, 以判定一統計量是否為充分統計量。只要是  $f(\mathbf{x}|\theta)$  可寫成一  $\mathbf{x}$  的函數, 乘上一  $T$  與  $\theta$  的函數, 即可知  $T$  為  $\theta$  之一充分統計量。通常並不難看出  $T$  為何, 見下例。

**例2.9** 設  $X$  有  $\mathcal{N}(0, \sigma^2)$  分佈,  $\sigma > 0$ , 即  $X$  之 p.d.f. 為

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

由分解定理,  $T_1(X) = X$ , 及  $T_2(X) = X^2$  皆為  $\sigma^2$  之充分統計量。從資料減縮的觀點, 我們會較偏好  $T_2(X)$ 。又  $T_3(X) = |X|$  亦為  $\sigma^2$  之一充分統計量, 且與  $T_2(X)$  等價。此三統計量當然也都是  $\sigma$  之充分統計量。

諸如  $X^4$ , 及  $e^{X^2}$  等, 皆與  $T_2(X)$  等價, 故皆為  $\sigma^2$  之充分統計量。

又設  $X_1, \dots, X_n$  為一組由  $\mathcal{N}(0, \sigma^2)$  分佈所產生之隨機樣本,  $\sigma > 0$ 。對  $n \geq 3$ , 由分解定理, 下述四統計量皆為  $\sigma^2$  之充分統計量。

$$T_4(\mathbf{X}) = (X_1, \dots, X_n),$$

$$T_5(\mathbf{X}) = (X_1^2, \dots, X_n^2),$$

$$T_6(\mathbf{X}) = (X_1^2 + \dots + X_m^2, X_{m+1}^2 + \dots + X_n^2), 1 \leq m < n,$$

$$T_7(\mathbf{X}) = X_1^2 + \dots + X_n^2。$$

又可看出對上述4個  $T_i$ ,  $i$  愈大,  $T_i$  愈有資料減縮的效果。

**例2.10** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{N}(\mu, \sigma^2)$  分佈所產生之隨機樣本,  $\mu \in R$ ,  $\sigma > 0$ ,  $n \geq 2$ 。則

$$(2.2) \quad f(\mathbf{x}|\mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\sum_{i=1}^n (x_i - \mu)^2 / (2\sigma^2)\right\}。$$

(i) 設  $\theta = (\mu, \sigma^2)$ 。則

$$(2.3) \quad f(\mathbf{x}|\mu, \sigma^2)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{-\sum_{i=1}^n x_i^2/(2\sigma^2) + \sum_{i=1}^n x_i \cdot \mu/\sigma^2 - n\mu^2/(2\sigma^2)\right\},$$

故由分解定理,  $T_1(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$  為  $\theta$  之一充分統計量。又  $f(\mathbf{x}|\mu, \sigma^2)$  可改寫為

$$(2.4) \quad f(\mathbf{x}|\mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\{-n(\bar{x}_n - \mu)^2/(2\sigma^2) - (n-1)s_n^2/(2\sigma^2)\},$$

其中  $\bar{x}_n = \sum_{i=1}^n x_i/n$ ,  $s_n^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2/(n-1)$ 。故亦得  $T_2(\mathbf{X}) = (\bar{X}_n, S_n^2)$  為  $\theta$  之一充分統計量。事實上  $(\bar{X}_n, S_n^2)$  與  $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$  等價(習題第6題)。

(ii) 設  $\theta = \mu$ , 即  $\sigma^2$  設為已知。則由(2.3)式,  $T_3(\mathbf{X}) = \sum_{i=1}^n X_i$  為  $\mu$  之一充分統計量。

(iii) 設  $\theta = \sigma^2$ , 即  $\mu$  設為已知。則由(2.3)式, 可得  $T_4(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$  為  $\sigma^2$  之一充分統計量。但若由(2.2)式, 得  $T_5(\mathbf{X}) = \sum_{i=1}^n (X_i - \mu)^2$  為  $\sigma^2$  之一充分統計量。而此  $T_5$  為  $T_4$  之一函數, 且  $T_4$  為二維,  $T_5$  為一維, 二者不等價, 故  $T_5$  較  $T_4$  更能減縮資料。

本例顯示, 使用分解定理, 有時得到表面上看起來不同, 其實為等價的統計量(如  $T_1$  與  $T_2$ )。有時得到的統計量尚可再減縮(如由  $T_4$  至  $T_5$ )。又由分解定理當然也都可得到  $(X_1, \dots, X_n)$  及  $(X_{(1)}, \dots, X_{(n)})$  為二組充分統計量。只是通常我們會找更能減縮資料的充分統計量。

**例2.11** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{U}(\alpha, \beta)$  分佈所產生之隨機樣本,  $\alpha < \beta$ 。則

$$\begin{aligned} f(\mathbf{x}|\alpha, \beta) &= (\beta - \alpha)^{-n} \prod_{i=1}^n I_{\{\alpha < x_i < \beta\}} \\ &= (\beta - \alpha)^{-n} I_{\{\alpha < x_{(1)}\}} I_{\{x_{(n)} < \beta\}}。 \end{aligned}$$

其中  $x_{(1)} = \min\{x_1, \dots, x_n\}$ ,  $x_{(n)} = \max\{x_1, \dots, x_n\}$ 。

(i) 若  $\alpha$  及  $\beta$  皆未知, 則  $(X_{(1)}, X_{(n)})$  為  $(\alpha, \beta)$  之一充分統計量;

- (ii) 若  $\beta$  已知, 則  $X_{(1)}$  為  $\alpha$  之一充分統計量;  
 (iii) 若  $\alpha$  已知, 則  $X_{(n)}$  為  $\beta$  之一充分統計量。

**註2.1** 如前, 對一  $R^n$  的子集合  $A$ ,  $n \geq 1$ , 令

$$I_A = \begin{cases} 1, & \text{若 } A \text{ 成立,} \\ 0, & \text{若 } A \text{ 不成立。} \end{cases}$$

例如, 取  $A = \{0 < x < y < 4\}$ , 則若  $x = 3, y = 2$ ,  $I_A$  便為 0。至於符號  $I_A(x)$  表若  $x \in A$ , 則  $I_A(x) = 1$ , 若  $x \notin A$ , 則  $I_A(x) = 0$ 。因此  $I_{\{\alpha < x_i < \beta\}} = I_{(\alpha, \beta)}(x_i)$ 。

**例2.12** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{U}[-\theta, \theta]$  分佈所產生之隨機樣本,  $\theta > 0$ 。仿例2.11, 可得  $(X_{(1)}, X_{(n)})$  為  $\theta$  之一充分統計量。但此顯然非最小充分統計量。事實上,  $(X_{(1)}, X_{(n)})$  的函數  $T(X_{(1)}, X_{(n)}) = \max\{|X_{(1)}|, |X_{(n)}|\}$  亦為充分統計量:

$$\begin{aligned} f(\mathbf{x}|\theta) &= (2\theta)^{-n} I_{\{-\theta \leq x_{(1)} \leq x_{(n)} \leq \theta\}} \\ &= (2\theta)^{-n} I_{\{\max\{|x_{(1)}|, |x_{(n)}|\} \leq \theta\}} \circ \end{aligned}$$

其中最後一等號成立之證明留在習題中(第18題)。最後, 利用定理2.2, 立即得  $\max\{|X_{(1)}|, |X_{(n)}|\}$  為  $\theta$  之最小充分統計量(亦見習題第18題)。

**註2.2** 在例2.12, 我們給更一般的結果: 設  $X_1, \dots, X_n$  為一組由如下之 p.d.f. 所產生之隨機樣本:

$$(2.5) \quad f(\mathbf{x}|\theta) = c(\theta)h(\mathbf{x})I_{\{q_1(\theta) \leq \mathbf{x} \leq q_2(\theta)\}} \circ$$

若  $q_1(\theta)$  為漸減, 且  $q_2(\theta)$  為漸增, 則  $T_1(\mathbf{X}) = \max\{q_1^{-1}(X_{(1)}), q_2^{-1}(X_{(n)})\}$  為  $\theta$  之一充分統計量; 若  $q_1(\theta)$  為漸增, 且  $q_2(\theta)$  為漸減, 則  $T_2(\mathbf{X}) = \min\{q_1^{-1}(X_{(1)}), q_2^{-1}(X_{(n)})\}$  為  $\theta$  之一充分統計量。又若(2.5)式中之指示函數改為  $I_{\{q_1(\theta) < \mathbf{x} < q_2(\theta)\}}$ , 結果仍相同。

如何利用定理2.2以檢驗例2.5中之 $T$ 為 $\theta$ 之一充分統計量呢？表面上看起來不太容易，其實還是可以的，見下例。

**例2.13** 承例2.5。首先 $X$ 之p.d.f.可改寫為

$$f(x|\theta) = \begin{cases} (1-\theta)/4 + \theta/4, & x = 1, 2, \\ (1-\theta)/4 + \theta/2, & x = 3, \\ (1-\theta)/4, & x = 4, \end{cases}$$

現考慮

$$T_1(X) = I_{\{1,2\}}(X) + 2I_{\{3\}}(X)。$$

且取

$$(2.6) \quad g(t|\theta) = \frac{1-\theta}{4} + \frac{\theta}{4}t, \quad h(x) = I_{\{1,2,3,4\}}(x),$$

則得

$$f(x|\theta) = g(T_1(x)|\theta)h(x),$$

由分解定理，即得 $T_1$ 為 $\theta$ 之一充分統計量。又例2.5中之 $T$ 與上述 $T_1$ 等價，故得 $T$ 亦為 $\theta$ 之一充分統計量。

有一類分佈可很容易地找到充分統計量，見下定理。證明並不難，留給讀者自行完成。

**定理2.3** 令 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta)$ 所產生之隨機樣本。假設 $f(x|\theta)$ 有下述形式：

$$(2.7) \quad f(x|\theta) = h(x)c(\theta)\exp\left\{\sum_{j=1}^k w_j(\theta)t_j(x)\right\}I_A(x),$$

其中 $h(x) \geq 0, c(\theta) \geq 0$ ，且 $I_A$ 不能與 $\theta$ 有關。則

$$T(\mathbf{X}) = \left(\sum_{i=1}^n t_1(X_i), \dots, \sum_{i=1}^n t_k(X_i)\right)$$

為 $\theta$ 之一充分統計量。

具有形式如(2.7)式所給之p.d.f.的分佈，便稱為屬於 $k$ 個參數之指數族(exponential family)。不少常見的分佈皆屬於指數族。

**例2.14**  $\mathcal{N}(\mu, \sigma^2)$ 分佈之p.d.f.可寫為

$$f(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-\mu^2/(2\sigma^2)} \exp\{(\mu/\sigma^2)x - x^2/(2\sigma^2)\}。$$

令 $\theta = (\mu, \sigma^2)$ 。對應(2.7)式， $h(x) = (2\pi\sigma^2)^{-1/2}$ ， $c(\theta) = e^{-\mu^2/(2\sigma^2)}$ ， $w_1(\theta) = \mu/\sigma^2$ ， $w_2(\theta) = -1/(2\sigma^2)$ ， $t_1(x) = x$ ， $t_2(x) = x^2$ 。因此若 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本，由定理2.3， $T(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 為 $\theta$ 之一充分統計量。由於 $T_1(\mathbf{X}) = (\bar{X}_n, S_n^2)$ 與 $T(\mathbf{X})$ 等價，故得 $T_1(\mathbf{X})$ 亦為 $\theta$ 之一充分統計量。

與例2.10比較，對指數族分佈，可同時採用分解定理或定理2.3以找出充分統計量。只是前者須有 $X_1, \dots, X_n$ 之聯合p.d.f.，而後者只須有一個 $X$ 之p.d.f.。本質上作法則並無太大差別。

至於是否有非指數族的分佈呢？也是有的，見下例。

**例2.15**  $\mathcal{H}(N, D, n)$ 分佈之p.d.f.為

$$f(x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} I_{\{\max\{0, n-N+D\} \leq x \leq \min\{n, D\}\}}。$$

不論參數為 $N, D, n, (N, n), (N, D), (D, n)$ ，或 $(N, D, n)$ ，明顯地，此分佈皆非指數族。

指數族之一特性是其p.d.f.為正之處，不能仰賴參數 $\theta$ 。在(2.7)式中，p.d.f.為0之處，就恰是 $h(x) = 0$ 之處。所以如 $\mathcal{U}(\alpha, \beta)$ 分佈便非指數族。二項分佈 $\mathcal{B}(m, p)$ ，若 $m, p$ 皆為未知，則非指數族(p.d.f.為正之處與 $m$ 有關)，但當 $\theta = p$ 時( $m$ 設為已知)為指數族。此由下述p.d.f.便立即得知：

$$\begin{aligned} f(x|p) &= \binom{m}{x} p^x (1-p)^{m-x} \\ &= \binom{m}{x} \exp\{x \log p + (m-x) \log(1-p)\}, x = 0, 1, \dots, m。 \end{aligned}$$

與(2.7)式相比, 得知當有樣本 $X_1, \dots, X_n$ 時,

$$T_1(\mathbf{X}) = \left( \sum_{i=1}^n X_i, \sum_{i=1}^n (m - X_i) \right)$$

為 $p$ 之一充分統計量。如果回到例2.7,  $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 便已是 $p$ 之充分統計量了, 為何此處 $T_1$ 較複雜? 不難看出 $T$ 與 $T_1$ 等價。事實上, 若知 $\sum_{i=1}^n X_i$ , 則 $\sum_{i=1}^n (m - X_i) = mn - \sum_{i=1}^n X_i$ 當然是多餘了。如果將 $f(x|p)$ 改寫為

$$\begin{aligned} f(x|p) &= \binom{m}{x} (1-p)^m (p/(1-p))^x \\ &= \binom{m}{x} (1-p)^m \exp\{x \log(p/(1-p))\}, \end{aligned}$$

便可寫出充分統計量 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 了。欲藉助定理2.3, 對 $f(x|\theta)$ 不同的寫法, 可能導致不同的充分統計量。但它們之間是等價的。

欲由定義2.2尋找最小充分統計量當然不是很實際, 幸好我們有下述定理, 其證明可見Casella and Berger (1990) pp.255-256。

**定理2.4** 設 $\mathbf{X} = (X_1, \dots, X_n)$ 之聯合p.d.f.為 $f(\mathbf{x}|\theta)$ 。又設存在一函數 $T(\mathbf{X})$ , 使得對任二樣本點 $\mathbf{x}$ 及 $\mathbf{y}$ ,  $f(\mathbf{x}|\theta)/f(\mathbf{y}|\theta)$ 與 $\theta$ 無關, 若且唯若 $T(\mathbf{x}) = T(\mathbf{y})$ 。則 $T(\mathbf{X})$ 為 $\theta$ 之一最小充分統計量。

**例2.16** 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。

(i) 設 $\theta = (\mu, \sigma^2)$ 。則

$$\begin{aligned} (2.8) \quad & \frac{f(\mathbf{x}|\mu, \sigma^2)}{f(\mathbf{y}|\mu, \sigma^2)} \\ &= \frac{\exp\{-\sum_{i=1}^n x_i^2/(2\sigma^2) + \sum_{i=1}^n x_i \cdot \mu/\sigma^2\}}{\exp\{-\sum_{i=1}^n y_i^2/(2\sigma^2) + \sum_{i=1}^n y_i \cdot \mu/\sigma^2\}} \\ &= \exp\left\{\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right) \cdot \mu/\sigma^2 - \left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right)/(2\sigma^2)\right\}. \end{aligned}$$

可看出此比值與 $\theta$ 無關, 若且唯若 $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ , 且 $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$ 。故由定理2.4,  $T_1(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 為 $\theta$ 之一最小充分統計量。

(ii) 設  $\theta = \mu$ ,  $\sigma^2$  則設為已知。由(2.8)式  $f(\mathbf{x}|\mu)/f(\mathbf{y}|\mu)$  要與  $\mu$  無關, 若且唯若  $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ 。故  $T_3(\mathbf{X}) = \sum_{i=1}^n X_i$  為  $\mu$  之一最小充分統計量。

(iii) 設  $\theta = \sigma^2$ ,  $\mu$  則設為已知。由(2.2)式得

$$(2.9) \quad \frac{f(\mathbf{x}|\sigma^2)}{f(\mathbf{y}|\sigma^2)} = \exp\left\{-\left(\sum_{i=1}^n (x_i - \mu)^2 - \sum_{i=1}^n (y_i - \mu)^2\right)/(2\sigma^2)\right\}。$$

故  $T_5(\mathbf{X}) = \sum_{i=1}^n (X_i - \mu)^2$  為  $\sigma^2$  之一最小充分統計量。但若由(2.8)式, 可能會誤以為  $T_4(\mathbf{X}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$  為  $\sigma^2$  之一最小充分統計量。顯然利用定理2.4時仍要小心為之。

**例2.17** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{B}(m, p)$  分佈所產生之隨機樣本。先設  $\theta = (m, p)$ 。則  $\mathbf{X} = (X_1, \dots, X_n)$  之聯合p.d.f. 為

$$f(\mathbf{x}|\theta) = \prod_{i=1}^n \binom{m}{x_i} p^{\sum_{i=1}^n x_i} (1-p)^{mn - \sum_{i=1}^n x_i}。$$

故

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\prod_{i=1}^n \binom{m}{x_i}}{\prod_{i=1}^n \binom{m}{y_i}} \left(\frac{p}{1-p}\right)^{\sum_{i=1}^n x_i - \sum_{i=1}^n y_i}$$

與  $\theta$  無關, 若且唯若  $x_{(1)} = y_{(1)}, \dots, x_{(n)} = y_{(n)}$ 。故  $T_1(\mathbf{X}) = (X_{(1)}, \dots, X_{(n)})$  為  $\theta$  之一最小充分統計量。若  $p$  已知, 仍得  $T_1(\mathbf{X})$  為  $m$  之一最小充分統計量。又若  $m$  已知, 則  $T_2(\mathbf{X}) = \sum_{i=1}^n X_i$  為  $p$  之一最小充分統計量。

**例2.18** 設  $X_1, \dots, X_n$  為一組由  $\mathcal{U}(\theta, \theta + 1)$  分佈所產生之隨機樣本,  $\theta \in R$ ,  $n \geq 2$ 。則  $\mathbf{X} = (X_1, \dots, X_n)$  之聯合p.d.f. 為

$$f(\mathbf{x}|\theta) = I_{\{x_{(n)} - 1 < \theta < x_{(1)}\}}。$$

故  $f(\mathbf{x}|\theta)/f(\mathbf{y}|\theta)$  與  $\theta$  無關, 若且唯若  $x_{(1)} = y_{(1)}, x_{(n)} = y_{(n)}$ , 且此時比值為1。再由定理2.4知,  $T(\mathbf{X}) = (X_{(1)}, X_{(n)})$  為  $\theta$  之一最小充分統計量。另外, 等價統計量  $(X_{(n)} - X_{(1)}, (X_{(n)} + X_{(1)})/2)$  亦為  $\theta$  之一最小充分統計量。

最小充分統計量的維數常與參數的維數相同(如例2.16)。但也不一定,可能大於(如例2.17及2.18),也可能小於。例如,若 $X$ 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈,則僅是一維的 $X$ 為 $(\mu, \sigma^2)$ 之一充分統計量。習題第21題及15題,亦分別為最小充分統計量的維數大於及小於參數的維數之例。

如果分佈屬於指數族,我們當然可很快地找出充分統計量;如果分解定理行得通,充分統計量亦可輕易地找出。但即使看不出如何分解,也不能立即得到結論說一統計量不為充分統計量。Sampson and Spencer(1976)提出幾個判別統計量不為充分,以及充分統計量不為最小的方法。我們列成底下三個定理。首先對隨機向量 $\mathbf{X} = (X_1, \dots, X_n)$ 之p.d.f. $f(\mathbf{x}|\theta)$ , 令

$$(2.10) \quad SPf = \{\mathbf{x} | f(\mathbf{x}|\theta) > 0, \text{對某}\theta\text{成立}\}。$$

$SPf$ 表以 $f$ 為p.d.f.之隨機向量(如果 $n = 1$ 則為隨機變數)取正值之處(即其支撐)。對 $\mathcal{N}(\mu, \sigma^2)$ 分佈,  $\theta = (\mu, \sigma^2)$ ,  $SPf = R$ ; 對 $\mathcal{P}(\theta)$ 分佈,  $\theta > 0$ ,  $SPf = \{0, 1, \dots\}$ ; 對 $\mathcal{U}(0, \theta)$ 分佈,  $\theta > 0$ ,  $SPf = (0, \infty)$ 。

**定理2.5** 設有一統計量 $T(\mathbf{X})$ 。若存在 $\theta_1, \theta_2$ , 及 $\mathbf{x}, \mathbf{y} \in SPf$ , 使得

$$(i) \quad T(\mathbf{x}) = T(\mathbf{y}),$$

$$(ii) \quad f(\mathbf{x}|\theta_1)f(\mathbf{y}|\theta_2) \neq f(\mathbf{x}|\theta_2)f(\mathbf{y}|\theta_1),$$

則 $T(\mathbf{X})$ 不為 $\theta$ 之一充分統計量。

**定理2.6** 設 $S(\mathbf{X})$ 為一最小充分統計量,  $T(\mathbf{X})$ 為一統計量。若存在 $\mathbf{x}, \mathbf{y} \in SPf$ , 使得 $T(\mathbf{x}) = T(\mathbf{y})$ , 且 $S(\mathbf{x}) \neq S(\mathbf{y})$ , 則 $T(\mathbf{X})$ 不為一充分統計量。

**定理2.7** 設 $S(\mathbf{X}), T(\mathbf{X})$ 為二充分統計量。若存在 $\mathbf{x}, \mathbf{y} \in SPf$ , 使得 $S(\mathbf{x}) = S(\mathbf{y})$ , 且 $T(\mathbf{x}) \neq T(\mathbf{y})$ , 則 $T(\mathbf{X})$ 不為一最小充分統計量。

此三定理之證明均不難, 利用定義2.1, 2.2及定理2.1即可證出。留給讀者練習自行證明(習題第28-30題)。

**例2.19** 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta)$ ,  $\theta > 0$ , 所產生之隨機樣本, 其

中

$$f(x|\theta) = \begin{cases} x/\theta^2 & , 0 < x \leq \theta, \\ (2\theta - x)/\theta^2 & , \theta < x < 2\theta, \\ 0 & , \text{其他。} \end{cases}$$

$f$ 之圖形與 $x$ 軸形成一由 $x = 0$ 至 $x = 2\theta$ 間之一三角形，圖形之最高點發生在 $x = \theta$ 。設 $n \geq 5$ ，底下我們看，如何利用定理2.5，以證明 $T(\mathbf{X}) = (\text{med}X, \max X)$ 並非 $\theta$ 之一充分統計量。

取 $\theta_1 = 9/8$ ,  $\theta_2 = 1$ ,  $\mathbf{x} = (1/4, 1/2, 5/4, 3/2, 1, 1, \dots, 1)$ ,  $\mathbf{y} = (1/2, 3/2, 1, 1, \dots, 1)$ 。則 $\text{med}x = \text{med}y = 1$ ,  $\max x = \max y = 3/2$ 。故 $T(\mathbf{x}) = T(\mathbf{y})$ 。因

$$\begin{aligned} f(\mathbf{x}|\theta_1) &= \frac{1}{4} \frac{1}{2} (2 \cdot \frac{9}{8} - \frac{5}{4}) (2 \cdot \frac{9}{8} - \frac{3}{2}) (\frac{9}{8})^{-2n} = \frac{3}{2^5} (\frac{9}{8})^{-2n}, \\ f(\mathbf{x}|\theta_2) &= \frac{1}{4} \frac{1}{2} (2 - \frac{5}{4}) (2 - \frac{3}{2}) = \frac{3}{2^6}, \\ f(\mathbf{y}|\theta_1) &= \frac{1}{2} (2 \cdot \frac{9}{8} - \frac{3}{2}) (\frac{9}{8})^{-2n} = \frac{3}{2^3} (\frac{9}{8})^{-2n}, \\ f(\mathbf{y}|\theta_2) &= \frac{1}{2} (2 - \frac{3}{2}) = \frac{1}{2^2}, \end{aligned}$$

故

$$\begin{aligned} f(\mathbf{x}|\theta_1)f(\mathbf{y}|\theta_2) &= \frac{3}{2^7} (\frac{9}{8})^{-2n}, \\ f(\mathbf{x}|\theta_2)f(\mathbf{y}|\theta_1) &= \frac{9}{2^9} (\frac{9}{8})^{-2n}, \end{aligned}$$

兩者不等。因此由定理2.5知,  $T(\mathbf{X}) = (\text{med}X, \max X)$ 不為 $\theta$ 之一充分統計量。

最後，以**分割**(partition)的概念，引進充分統計量，乃是被不少人認為是一很好的理解充分統計量之意義的方式。可參考Lindgren(1976)一書，在此略過。