

第五章

資料減縮

5.1 前言

前面幾章我們主要討論機率論，自本章開始要討論一些關於統計的題材。

機率論與統計學裡所探討的問題是不太一樣的。在機率論裡，可以假設 $X_1, \dots, X_n$ 為 i.i.d. 以 $\mathcal{N}(1, 2)$ 為共同分佈。然後求諸如機率 $P(\sum_{i=1}^n X_i > 3)$ ，或 $\sum_{i=1}^k (X_i - 1)^2 / \sum_{i=k+1}^n (X_i - 1)^2$ 之分佈，其中 $1 \leq k \leq n-1$ 。在這些過程中，一直是將 $X_1, \dots, X_n$ 當做已知分佈的隨機變數。當然此分佈的產生也可依過去的經驗，或所涉及之隨機現象的性質。如人的智商、身高等，常有常態分佈；另外，若相信一銅板為公正，則投擲10次後，所得正面數便有 $B(10, 1/2)$ 分佈等。

統計學裡，則是由觀測到的資料(或說數據、樣本)，以對模型做一些推估。例如，就算相信人的身高可以常態分佈當模型，但其中的參數為何？或即使認為每次投擲銅板的結果為 i.i.d.，但銅板出現正面的機率是否真如一般人以為的 $1/2$ ？由於是隨機的現象，所以每次的觀測是不盡相同的。例如，設一盒子中放有8個白球，2個紅球(但未透露)。如果隨機地取10個球(每次取出後放回)皆得到紅球，則很可能會判定盒中是紅球

多。這個判定算是合理的。機率理論(如大數法則)告訴我們,若取樣夠多,所得紅球數之相對頻率,接近0.2的機率會很大。但卻不是百分之百保證有多接近0.2。所以錯誤的判定,在統計學裡是無法避免的。而這並非統計的錯,判斷不會有錯誤,只有非隨機現象。統計學裡要追求的,便是要發展好的統計方法,以有效地做決策,並減少誤判的機率。

如前所述,一旦我們建立(或接受)了一統計模式,便想對其中未知的參數做一些推估,於是我們會重複地觀察,因而得到樣本 $X_1, \dots, X_n$ 。如果 $X_1, \dots, X_n$ 為獨立,並以函數 f 為其共同的p.d.f.,則我們便有一由母體 f 所產生之樣本數為 n 之隨機樣本。一般所謂“隨機樣本”,便表這些樣本為i.i.d.(見上一章定義1.1)。

我們常說收集資料(或說數據, data),現在就是有了資料 $X_1, \dots, X_n$ 。利用這些資料要估計某未知的參數 θ (θ 也可以是一向量)。估計要準確,樣本數 n 往往便不會太小。但如此一來資料便有不少。如何對資料作一些減縮呢?這是本章主要想討論的。

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 θ 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

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

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 $\bar{X}_n + \theta$ 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 θ 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 θ 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 θ 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 θ , 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 θ 提供所有資訊。

首先有

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

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 θ 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

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

前者與 θ 無關, 所以未能提供對 θ 之任何資訊, 而後者包含資料中所有對 θ 之資訊。所以若要對 θ 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 θ 有關。對一統計量 $T(\mathbf{X})$, 若給定 $T(\mathbf{X})$, $\mathbf{X}$ 之條件分佈“與 θ 無關”(does not depend on θ), 則 $T(\mathbf{X})$ 稱為 θ 之一充分統計量。

若 T 為 θ 之一充分統計量, 則知道 T 之後, 便掌握對 θ 之所有資訊(即知道 T 之後, $\mathbf{X}$ 之變化已不再仰賴 θ), 也就是樣本中已不再含有更多關於 θ 之資訊。知道樣本的其他資訊, 對 θ 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 θ), 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 θ 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 θ 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 θ 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 θ 有關, 統計量 $T = (X_1, \dots, X_n)$, 必為 θ 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 θ 無關。由於充分統計量是為減縮資料(但不減少對參數 θ 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 θ 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 θ 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 θ 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 σ^2 之充分統計量, 其中 T_2 與 T_3 等價。 T_2 與 T_3 皆爲 σ^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 θ 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 θ 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 θ 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 θ 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 θ 無關, 則在給定 $T = t$ 之下, $T, Y_2, \dots, Y_n$ 之條件分佈也與 θ 無關。因此在給定 $T = t$ 之下, $X_1, \dots, X_n$ 之條件分佈與 θ 無關。由是便得證 T 為 θ 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 θ 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 θ 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 θ ,

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

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 θ 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 θ 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 θ 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 θ 的函數, 即可知 T 為 θ 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 σ^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 σ^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 σ 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 σ^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 σ^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 θ 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 θ 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 σ^2 設為已知。則由(2.3)式, $T_3(\mathbf{X}) = \sum_{i=1}^n X_i$ 為 μ 之一充分統計量。

(iii) 設 $\theta = \sigma^2$, 即 μ 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 σ^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 σ^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 α 及 β 皆未知, 則 $(X_{(1)}, X_{(n)})$ 為 (α, β) 之一充分統計量;

- (ii) 若 β 已知, 則 $X_{(1)}$ 為 α 之一充分統計量;
 (iii) 若 α 已知, 則 $X_{(n)}$ 為 β 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 θ 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 θ 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 θ 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 θ 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 爲 θ 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.可改寫爲

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

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 爲 θ 之一充分統計量。又例2.5中之 T 與上述 T_1 等價，故得 T 亦爲 θ 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 $X_1, \dots, X_n$ 爲一組由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 θ 有關。則

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

爲 θ 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 θ 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 θ 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 θ 。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 θ 無關, 若且唯若 $T(\mathbf{x}) = T(\mathbf{y})$ 。則 $T(\mathbf{X})$ 為 θ 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 θ 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 θ 之一最小充分統計量。

(ii) 設 $\theta = \mu$, σ^2 則設為已知。由 (2.8) 式 $f(\mathbf{x}|\mu)/f(\mathbf{y}|\mu)$ 要與 μ 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 μ 之一最小充分統計量。

(iii) 設 $\theta = \sigma^2$, μ 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 σ^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 σ^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 θ 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 θ 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 θ 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 θ 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 θ 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 (μ, σ^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 θ_1, θ_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 θ 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 θ 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 θ 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。

5.3 其他統計量

上一節我們討論充分統計量，這種統計量可說是包含樣本中所有關於參數 θ 之資訊。本節我們討論一些在統計推論中，也相當有用的統計量。先給下述定義。

定義3.1 一統計量 $S(\mathbf{X})$ ，若其分佈與參數 θ 無關，便稱為 $(\theta\text{-})$ 一輔助統計量(ancillary statistic)。

表面上看來，輔助統計量並未包含關於參數 θ 之資訊。但有時一輔助統計量與其他統計量結合後，卻能包含一些對 θ 有用的資訊。先給一個例子。

例3.1 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， $-\infty < \theta < \infty$ 。令 $X_{(1)} < \dots < X_{(n)}$ 表順序統計量。底下我們驗證全距統計量 $R = X_{(n)} - X_{(1)}$ 為一輔助統計量。

首先 $(X_{(1)}, X_{(n)})$ 之聯合p.d.f.為

$$f_{X_{(1)}, X_{(n)}}(y_1, y_n | \theta) = n(n-1)(y_n - y_1)^{n-2}, \theta < y_1 < y_n < \theta + 1。$$

再利用變數代換，令 $R = X_{(n)} - X_{(1)}$ ， $V = (X_{(1)} + X_{(n)})/2$ ，得 R, V 之聯合p.d.f.為

$$f_{R,V}(r, v | \theta) = n(n-1)r^{n-2}, 0 < r < 1, \theta + r/2 < v < \theta + 1 - r/2。$$

由此得 R 之邊際p.d.f.為

$$\begin{aligned} h_R(r | \theta) &= \int_{\theta+r/2}^{\theta+1-r/2} n(n-1)r^{n-2} dv \\ &= n(n-1)r^{n-2}(1-r), 0 < r < 1。 \end{aligned}$$

故 R 有 $\mathcal{Be}(n-1, 2)$ 分佈(可與第四章例5.4比較)。由於 R 的分佈與 θ 無關，所以 R 為一輔助統計量。

在上例中， X_i 's 的分佈為位置族(location family)，所以 R 的分佈與 θ 無關是當然的。在此若一隨機變數 X 之p.d.f.有 $f(x - \mu)$ 的形式，其中 f 為

一p.d.f., μ 為一常數, X 的分佈便稱屬於**位置族**, 而 μ 稱為**位置參數**。例如, 若 X 有 $\mathcal{N}(\mu, 1)$ 分佈, 則 X 屬於位置參數為 μ 之位置族。若 X 之p.d.f. 為 $f(x) = \lambda e^{-\lambda(x-\theta)}, x > \theta$, 亦為位置族, 位置參數為 θ 。通常以一分佈當源頭, 經變數平移的變換, 便得到一位置族分佈。如若 X 有 $\mathcal{U}(0, 1)$ 分佈, 令 $Y = X + \theta, \theta > 0$, 便得 $\mathcal{U}(\theta, \theta + 1)$ 分佈族。所謂“族”, 是指有一集合的分佈, 以某一參數為指標, 變動參數, 便得到不同的分佈。如 $\{f(x - \mu), \mu \in R\}$ 即為一分佈族。

底下我們給一例3.1之推廣。

例3.2 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x - \theta)$ 所產生之隨機樣本, $-\infty < x < \infty$ 。則 $R = X_{(n)} - X_{(1)}$ 為一輔助統計量:

令 $Z_i = X_i - \theta, i = 1, \dots, n$, 則 $Z_1, \dots, Z_n$ 為i.i.d., 以 $f(x)$ 為其共同p.d.f.。又

$$\begin{aligned} F_R(r|\theta) &= P(R \leq r) \\ &= P(X_{(n)} - X_{(1)} \leq r) \\ &= P(\max(Z + \theta) - \min(Z + \theta) \leq r) \\ &= P(\max Z - \min Z + \theta - \theta \leq r) \\ &= P(Z_{(n)} - Z_{(1)} \leq r)。 \end{aligned}$$

由於 $Z_1, \dots, Z_n$ 之分佈與 θ 無關, 所以 $Z_{(n)} - Z_{(1)}$ 之分佈亦與 θ 無關。故得證 R 為一輔助統計量。

事實上, 對位置族分佈, 可給出無限多個輔助統計量。諸如 $X_2 - X_1, 2X_3 + 3X_4 - X_1 - 4X_2$ 及 $(X_4 - X_1)(X_{(2)} - X_{(1)})$ 等, 皆為輔助統計量。

設 $f(x)$ 為一p.d.f., 若 X 之p.d.f. 為 $(1/\sigma)f(x/\sigma), \sigma > 0$, 便稱 X 屬於**尺度族**(scale family), 而 σ 稱為**尺度參數**。尺度族分佈很多。如在 $\mathcal{N}(\mu, \sigma^2)$ 分佈中, σ 為尺度參數; 在 $\Gamma(\alpha, \beta)$ 分佈中, β 為尺度參數。尺度族分佈, 亦可以一分佈當源頭, 經尺度的變換而得所有分佈。如若 X 有 $\mathcal{N}(0, 1)$ 分佈, 令 $Y = \sigma X, \sigma > 0$, 便得到一尺度族分佈。

例3.3 設 $X_1, \dots, X_n$ 為一組由 $p.d.f. f(x/\sigma)/\sigma$, $\sigma > 0$, 所產生之隨機樣本。令 $Z_i = X_i/\sigma, i = 1, \dots, n$, 則 $Z_1, \dots, Z_n$ 為 i.i.d. 以 $f(x)$ 為其共同 p.d.f., 且

$$\begin{aligned} & P\left(\frac{X_1}{X_n} \leq y_1, \dots, \frac{X_{n-1}}{X_n} \leq y_{n-1}\right) \\ &= P\left(\frac{\sigma Z_1}{\sigma Z_n} \leq y_1, \dots, \frac{\sigma Z_{n-1}}{\sigma Z_n} \leq y_{n-1}\right) \\ &= P\left(\frac{Z_1}{Z_n} \leq y_1, \dots, \frac{Z_{n-1}}{Z_n} \leq y_{n-1}\right). \end{aligned}$$

由於 $Z_1, \dots, Z_n$ 之分佈與 σ 無關, 所以 $Z_1/Z_n, Z_2/Z_n, \dots, Z_{n-1}/Z_n$ 之分佈亦與 σ 無關。故得證 $X_1/X_n, X_2/X_n, \dots, X_{n-1}/X_n$ 為一輔助統計量。

舉一例來看: 設 X_1, X_2 為 i.i.d 皆有 $\mathcal{N}(0, \sigma^2)$ 分佈, 由上述結果知, X_1/X_2 之分佈與 σ 無關。而當 $\sigma = 1$ 時, 可證明 X_1/X_2 有 $\mathcal{C}(0, 1)$ 分佈, 故對 $\forall \sigma > 0$, X_1/X_2 亦有 $\mathcal{C}(0, 1)$ 分佈。

另外, 當然諸如 $X_{(1)}/X_{(2)}$ 也是輔助統計量。

我們知道最小充分統計量是最能達到資料減縮, 且仍能保有關於參數 θ 的所有資訊之統計量。亦即最小充分統計量會消除樣本中與 θ 無關的資訊, 也就是把不需要的資訊刪除(資料減縮), 而只留下與 θ 有關的資訊。因輔助統計量與 θ 無關, 所以我們可能會猜測, 最小充分統計量與輔助統計量會獨立。可惜此一猜測並不正確。事實上, 在例2.18中, 對由 $\mathcal{U}(\theta, \theta + 1)$ 分佈所產生的隨機樣本 $X_1, \dots, X_n$, 我們已說明 $(X_{(n)} - X_{(1)}, (X_{(n)} + X_{(1)})/2)$ 為 θ 之一最小充分統計量。又在例3.1中, 我們曾驗證 $X_{(n)} - X_{(1)}$ 為 θ 之一輔助統計量。輔助統計量為最小充分統計量的一分量, 兩者當然不獨立。

雖然輔助統計量之分佈與參數 θ 無關, 但有時卻能對 θ 之推估, 提供重要的資訊。見下例。

例3.4 設 X_1, X_2 為由 $\mathcal{D}\text{-}\mathcal{U}[\theta, \theta + 2]$ 分佈所產生之隨機樣本, 其中 θ 為一整數。則 X_1, X_2 之聯合 p.d.f. 為

$$\begin{aligned} f(x_1, x_2|\theta) &= \frac{1}{9} I_{\{\theta \leq x_1, x_2 \leq \theta+2\}} \\ &= \frac{1}{9} I_{\{\theta \leq x_{(1)} \leq x_{(2)} \leq \theta+2\}}, \end{aligned}$$

其中 $x_{(1)}, x_{(2)}$ 為整數。可驗證 $(X_{(1)}, X_{(2)})$ 為 θ 之最小充分統計量。因此 (R, V) 亦為最小充分統計量，其中如前 $R = X_{(2)} - X_{(1)}$, $V = (X_{(1)} + X_{(2)})/2$ 。由於 X 有位置族分佈，由例3.2知， R 為一輔助統計量。

R 是一輔助統計量，會如何對 θ 提供資訊呢？假設有一樣本點 (r, v) ，其中 v 為一整數。如果只知道 v ，則 θ 有三個可能的值： $v, v-1, v-2$ （因 v 為整數，所以不會有 $X_{(1)} = \theta, X_{(2)} = \theta+1$ ，或 $X_{(1)} = \theta+1, X_{(2)} = \theta+2$ ）。若只知道 $V = v$ ，則三者皆為 θ 之可能的值。現若加上知道 $R = 2$ ，則必定是 $X_{(1)} = \theta, X_{(2)} = \theta+2$ ，所以 $\theta = v-1$ 。故知道輔助統計量 R 之值，會增進對 θ 之了解。當然，若只知輔助統計量 R 之值，對 θ 是無法提供任何資訊的。而即使只知 V 之值，仍對 θ 提供不少資訊（ θ 只剩三個可能的值）。

至於若 v 不為整數，則 $v = v_1 + 1/2$ ，其中 v_1 為整數，此時必有 $X_{(1)} = \theta, X_{(2)} = \theta+1$ ，或 $X_{(1)} = \theta+1, X_{(2)} = \theta+2$ ，且 R 必定為 1。故知道 R 之值，不會增進對 θ 之了解。

在不少常見的實例中，我們直觀上所認定最小充分統計量與輔助統計量獨立的想法倒是對的。我們先需要下述定義。

定義3.2 設 $T = T(\mathbf{X})$ 為一統計量， T 之 p.d.f. 為 $f(t|\theta), \theta \in \Omega$ 。對一函數 g ，若 $E_\theta(g(T)) = 0, \forall \theta \in \Omega$ ，便導致 $P_\theta(g(T) = 0) = 1, \forall \theta \in \Omega$ ，則稱 T 為**完備的**(complete)，或說 T 為一**完備統計量**(complete statistic)。此時也稱此分佈族為完備的。

在上述定義中，寫 E_θ, P_θ 是為了強調期望值及機率值與 θ 有關，有時會略去 θ 。須注意完備性是一分佈族的性質，而非針對一特定的分佈。例如，若 X 有 $\mathcal{N}(0, 1)$ 分佈，取 $g(x) = x$ ，則 $E(g(X)) = E(X) = 0$ ，但 $P(g(X) = 0) = P(X = 0) = 0 \neq 1$ 。不過此處只有一特定的分佈 $\mathcal{N}(0, 1)$ ，而非有一分佈族。若 X 有 $\mathcal{N}(\theta, 1)$ 分佈， $-\infty < \theta < \infty$ ，稍後我們會看

到 $T(X) = X$ 的確為一完備統計量。所以分佈族 $\mathcal{N}(\theta, 1)$, $-\infty < \theta < \infty$, 為完備的。又同一統計量對某一分佈族可能為完備統計量, 但對另一分佈族卻不為完備統計量。是否為完備統計量, 主要是看分佈族, 而不光是統計量本身。

例3.5 設 T 有 $B(n, \theta)$ 分佈, $0 < \theta < 1$ 。又設函數 g 滿足 $E_\theta(g(T)) = 0$ 。則

$$\begin{aligned} 0 = E_\theta(g(T)) &= \sum_{t=0}^n g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} \\ &= (1-\theta)^n \sum_{t=0}^n g(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t, \quad 0 < \theta < 1. \end{aligned}$$

由此得

$$(3.1) \quad \sum_{t=0}^n g(t) \binom{n}{t} r^t = 0,$$

其中

$$r = \frac{\theta}{1-\theta} = \frac{1}{1-\theta} - 1 > 0,$$

r 取值在 $(0, \infty)$ 。(3.1)式左側為一個 r 之 n 次多項式, 且對每一 $r > 0$ 皆使其為0。也就是方程式(3.1)之根的個數超過 n 。故(3.1)式左側多項式之係數須皆為0。由於 $\binom{n}{t} \neq 0$, 故 $g(t) = 0, t = 0, 1, \dots, n$ 。又由於 T 取值就在 $0, 1, \dots, n$, 故得 $g(T) = 0$ 。因此 T 為一完備統計量。

另外, 若將原假設改為 $\theta \in \{0.1, 0.2, 0.3\}$, 即 θ 只有3個可能的值, 且設 $n \geq 4$, 則 T 便非完備統計量了。所以一統計量是否為完備, 與參數取值的範圍有關。

例3.6 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{P}(\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。即 X_i 's 之共同p.d.f.為

$$f(x) = \frac{e^{-\theta} \theta^x}{x!}, x = 0, 1, 2, \dots。$$

由於 X_i 有指數族分佈, 故 $T = T(\mathbf{X}) = \sum_{i=1}^n X_i$ 為 θ 之一充分統計量。又 T 有 $\mathcal{P}(n\theta)$ 分佈。底下證明 T 為完備統計量。

假設函數 g 滿足 $E_\theta(g(T)) = 0, \forall \theta > 0$ 。即有

$$e^{-n\theta} \sum_{j=0}^{\infty} \frac{g(j)(n\theta)^j}{j!} = 0, \theta > 0。$$

由此又得

$$\sum_{j=0}^{\infty} \frac{g(j)n^j}{j!} \theta^j = 0, \theta > 0。$$

我們知道一個 θ 之冪級數，若在實數上一區間中恆為0，則其係數必皆為0。利用此結果，即得 $g(j) = 0, \forall j \geq 0$ 。證畢。

例3.7 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生之隨機樣本， $\theta > 0$ 。易見 $T = T(\mathbf{X}) = X_{(n)}$ 為 θ 之一充分統計量(見例2.11)。又 T 之p.d.f.為

$$f(t|\theta) = nt^{n-1}\theta^{-n}, 0 < t < \theta。$$

假設函數 g 滿足 $E_\theta(g(T)) = 0, \forall \theta > 0$ ，即

$$\int_0^\theta g(t)nt^{n-1}\theta^{-n}dt = 0, \theta > 0。$$

由此得

$$\int_0^\theta g(t)t^{n-1}dt = 0, \theta > 0。$$

由勒貝格(Lebesgue, 1875-1941)積分理論，上式導致對 $\forall \theta > 0, P_\theta(g(T) = 0) = 1$ (如果 g 為連續，由微積分基本定理得 $g(\theta)\theta^{n-1} = 0, \theta > 0$ ，因此 $g(\theta) = 0, \theta > 0$)。得證 T 為完備統計量。

例3.8 設 X 有 $\mathcal{H}(N, N\theta, n)$ 分佈， N, n 為已知， $\theta \in \{0, 1/N, 2/N, \dots, 1\}$ 。 X 當然為 θ 之一最小充分統計量。底下證明 X 為完備統計量，或者說此分佈族是完備的。

我們要證明若

$$E_\theta(g(X)) = 0, \theta = 0, \frac{1}{N}, \frac{2}{N}, \dots, 1,$$

則 $g(k) = 0, \forall k = 0, 1, \dots, n$ 。注意 n 要滿足 $n \leq N\theta$ 。首先

$$E_{\theta}(g(X)) = \sum_{k=0}^n \frac{\binom{N\theta}{k} \binom{N-N\theta}{n-k}}{\binom{N}{n}} g(k)。$$

因 $E_0(g(X)) = g(0)$, 故得 $g(0) = 0$ 。其次令 $\theta = 1/N$, 則由

$$E_{1/N}(g(X)) = \frac{\binom{1}{0} \binom{N-1}{n}}{\binom{N}{n}} g(0) + \frac{\binom{1}{1} \binom{N-1}{n-1}}{\binom{N}{n}} g(1) = 0,$$

立即得 $g(1) = 0$ 。藉由數學歸納法的討論, 使得 $g(0) = g(1) = \dots = g(n) = 0$ 。證畢。

利用完備性, 底下我們給一最小充分統計量與每一輔助統計量獨立的條件。此即**巴蘇定理**(Basu's theorem), 為古典統計學中一優美的結果。此定理將**充分性**、**輔助統計量**及**獨立性**等三個看起來似乎毫不相干的概念連接起來。

定理3.1 設 $T(\mathbf{X})$ 為一完備充分統計量, 則 $T(\mathbf{X})$ 與每一輔助統計量獨立。

證明. 我們只證明離散型分佈的情況。連續型分佈的證明類似。

設 $S(\mathbf{X})$ 為一輔助統計量, 則對 $\forall s \in R, P(S(\mathbf{X}) = s)$ 與 θ 無關。又因 $T(\mathbf{X})$ 為一充分統計量,

$$P(S(\mathbf{X}) = s | T(\mathbf{X}) = t) = P(\mathbf{X} \in \{\mathbf{x} : S(\mathbf{x}) = s\} | T(\mathbf{X}) = t)$$

與 θ 無關, 只要 $P(T(\mathbf{X}) = t) > 0$ 。欲證 $S(\mathbf{X})$ 與 $T(\mathbf{X})$ 獨立, 若能證出

$$(3.2) \quad P(S(\mathbf{X}) = s | T(\mathbf{X}) = t) = P(S(\mathbf{X}) = s)$$

對每個 $s, t \in R$, 只要 $P(T(\mathbf{X}) = t) > 0$, 皆成立即可。

現由總機率法則,

$$(3.3) \quad P(S(\mathbf{X}) = s) = \sum_t P(S(\mathbf{X}) = s | T(\mathbf{X}) = t) P_{\theta}(T(\mathbf{X}) = t),$$

其中和是對所有可能的 t 相加。又因 $\sum_t P_{\theta}(T(\mathbf{X}) = t) = 1$, 故

$$(3.4) \quad P(S(\mathbf{X}) = s) = \sum_t P(S(\mathbf{X}) = s) P_{\theta}(T(\mathbf{X}) = t)。$$

由(3.3)及(3.4)式得

$$(3.5) \sum_t (P(S(\mathbf{X}) = s | T(\mathbf{X}) = t) - P(S(\mathbf{X}) = s)) P_\theta(T(\mathbf{X}) = t) = 0。$$

若令 $g(t) = P(S(\mathbf{X}) = s | T(\mathbf{X}) = t) - P(S(\mathbf{X}) = s)$, 則(3.5)式可改寫為

$$E_\theta(g(T)) = \sum_t g(t) P_\theta(T(\mathbf{X}) = t) = 0, \forall \theta。$$

由於 $T(\mathbf{X})$ 為一完備統計量, 故得對每一 t , $g(t) = 0$ 。亦即(3.2)式對每一 t 成立。證畢。

有些書會在巴蘇定理中, 要求 T 為最小充分統計量。當然我們也看到在巴蘇定理的證明中, 並未用到充分統計量要為最小。事實上, 若最小充分統計量存在, 則一充分統計量, 只有當它是最小充分統計量時才是完備的。也就是在充分統計量中, 只有最小充分統計量才是完備的(見Lehmann(1983)p.46)。在定理的敘述中加進“最小”, 是為了提醒大家, 不是最小充分統計量就不用考慮了。

巴蘇定理可用來判別二統計量為獨立, 而不須求出其聯合分佈。欲使用此定理, 要先檢驗一統計量為完備, 有時這是分析學上一不容易的問題。幸好我們常考慮的分佈, 都是指數族。我們敘述一有用的結果如下, 證明則略去。

定理3.2 設 $X_1, \dots, X_n$ 為一組由 k 個參數之指數族所產生之隨機樣本, 即設 X 之共同 p.d.f. 如(2.7)式所給。又設 $C = (w_1(\theta), \dots, w_k(\theta))$ 之值域, 即集合 $\{w_1(\theta), \dots, w_k(\theta), \theta \in \Omega\}$ 有一非空的內部(即包含一 k 維開矩形)。則

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

為完備充分統計量。

例3.9 利用定理3.2, 可得對諸如 $\{\mathcal{B}(n, p), 0 < p < 1\}$, $\{\mathcal{P}(\lambda), \lambda > 0\}$ 分佈族, $\sum_{i=1}^n X_i$ 皆為完備充分統計量(分別見例3.5及3.6)。至於超幾何分佈及均勻分佈皆非指數族, 因此定理3.2不適用。

對 $\{\mathcal{N}(\mu, \sigma^2), \mu \in R, \sigma > 0\}$ 分佈族, 取 $w_1(\theta) = \mu/\sigma^2, w_2(\theta) = -1/(2\sigma^2)$, $t_1(x) = x, t_2(x) = x^2$, 而 $C = (w_1(\theta), w_2(\theta))$ 之值域為 R^2 。由定理3.2, 得 $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 為完備充分統計量, 因此等價統計量 $(\bar{X}_n, S_n^2)$ 亦為完備充分統計量。

對 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 設 σ^2 為已知, 令 $\theta = \mu$ 。此時只有 $w_1(\theta) = \mu/\sigma^2, t_1(x) = x$, 而 $C = (w_1(\theta))$ 之值域為實數集合 R , 由定理3.2, 得 $\sum_{i=1}^n X_i$ 為完備充分統計量。現設 μ 為已知, 令 $\theta = \sigma^2$ 。此時仍可取 $w_1(\theta) = \mu/\sigma^2, w_2(\theta) = -1/(2\sigma^2), t_1(x) = x, t_2(x) = x^2$, 但因 $C = (w_1(\theta), w_2(\theta)) = (\mu/\sigma^2, -1/(2\sigma^2))$ 之值域為 R^2 中一過原點之直線, 並未包含一個二維的開矩形, 故定理3.2不適用。也就是無法得到 $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 為完備充分統計量。

事實上, 若取 $g(x) = x - n\mu$, 則因

$$E_\theta(g(\sum_{i=1}^n X_i)) = E_\theta(\sum_{i=1}^n X_i - nu) = 0, \forall \theta > 0,$$

故得 $\sum_{i=1}^n X_i$ 不為完備統計量。因此 $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 不為完備充分統計量。但不難看出此時 $\sum_{i=1}^n (X_i - \mu)^2$ 為完備充分統計量。

對同一統計量(如 $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$), 因參數的不同 $((\mu, \sigma^2)$ 及 σ^2), 導致分佈族的不同, 有時為完備統計量, 有時則不是。此點應可理解。

最後對 $\{\Gamma(\alpha, \beta), \alpha, \beta > 0\}$ 分佈族, $(\sum_{i=1}^n X_i, \sum_{i=1}^n \log X_i)$, 或等價統計量 $(\sum_{i=1}^n X_i, \prod_{i=1}^n X_i)$, 為完備充分統計量。而若 α 為已知, 令 $\theta = \beta$, 則易見 $\sum_{i=1}^n X_i$ 為完備充分統計量; 若 β 為已知, 令 $\theta = \alpha$, 則 $\sum_{i=1}^n \log X_i$, 或等價統計量 $\prod_{i=1}^n X_i$, 為完備充分統計量。

例3.10 設 $X_1, \dots, X_n$ 為一組隨機樣本。令 $\bar{X}_n = \sum_{i=1}^n X_i/n$ 及 $S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2/(n-1)$, 分別表其樣本平均及樣本變異數。我們已證明若 X_i 's 之共同分佈為 $\mathcal{N}(\mu, \sigma^2)$, 則 $\bar{X}_n$ 與 S_n^2 獨立(見第四章定理3.1)。底下我們看如何利用巴蘇定理, 證明此一統計學中重要的結果。

首先固定 σ^2 。由例3.9知, $\bar{X}_n$ 為一完備統計量。又對一位置族分佈(注意 σ^2 固定, 而 μ 為位置參數), S_n^2 為一輔助統計量(其證明可仿例3.2之推

導)。再利用巴蘇定理, 即得證 S_n^2 與完備統計量 $\overline{X}_n$ 獨立。

雖是在 σ^2 固定之下證出 $\overline{X}_n$ 與 S_n^2 獨立, 但因 σ^2 為一常數, 知道 σ^2 後, 對 $\overline{X}_n$ 與 S_n^2 之聯合分佈並無任何影響。故得證對任意 $\mu, \sigma^2, \overline{X}_n$ 與 S_n^2 皆獨立。

例3.11 設 $X_1, \dots, X_n$ 相互獨立, 且對 $\forall i = 1, \dots, n$, X_i 有 $\Gamma(\alpha_i, \beta)$ 分佈, $\alpha_i, \beta > 0$ 。即 X_i 之p.d.f.為

$$f_i(x) = \frac{x^{\alpha_i-1}e^{-x/\beta}}{\Gamma(\alpha_i)\beta^{\alpha_i}} = \frac{1}{\beta} \frac{1}{\Gamma(\alpha_i)} \left(\frac{x}{\beta}\right)^{\alpha_i-1} e^{-x/\beta}, x > 0。$$

故 X_i 有尺度參數 β 。

先將 $\alpha_i, i = 1, \dots, n$, 視為已知, 且令 $T_i = \sum_{j=1}^i X_j, i = 1, \dots, n$ 。則 T_i 有 $\Gamma(\sum_{j=1}^i \alpha_j, \beta)$ 分佈。由例3.9知, 對 $X_1, \dots, X_i, T_i$ 為完備充分統計量。現令

$$Y_i = \frac{T_i}{T_{i+1}} = \frac{T_i}{T_i + X_{i+1}} = \frac{X_1 + \dots + X_i}{X_1 + \dots + X_{i+1}}, i = 1, \dots, n-1。$$

由於 X_i 's為尺度族, 仿例3.3之推導, 可得 $\mathbf{Y}_i = (Y_1, \dots, Y_i), i = 1, \dots, n-1$, 為輔助統計量。因此巴蘇定理導致對 $\forall i = 2, \dots, n, T_i$ 與 $\mathbf{Y}_{i-1}$ 獨立。又因 $(\mathbf{Y}_{i-1}, T_i)$ 為 $X_1, \dots, X_i$ 之函數, 故 $(\mathbf{Y}_{i-1}, T_i)$ 與 X_{i+1} 獨立, $\forall 2 \leq i \leq n-1$ 。由此即得 $\mathbf{Y}_{i-1}, T_i, X_{i+1}$ 相互獨立, $\forall 2 \leq i \leq n-1$ 。此結果又導致 $\mathbf{Y}_{i-1}$ 與

$$Y_i = \frac{T_i}{T_i + X_{i+1}}$$

獨立。亦即對 $\forall i = 2, \dots, n-1, (Y_1, \dots, Y_{i-1})$ 與 Y_i 獨立。又已指出 $\mathbf{Y}_{n-1} = (Y_1, \dots, Y_{n-1})$ 與 T_n 獨立。綜上所述, 即得 $Y_1, \dots, Y_{n-1}, T_n$ 相互獨立。換句話說, 對此組獨立有gamma分佈之隨機變數 $X_1, \dots, X_n$, 下述隨機變數相互獨立:

$$\frac{X_1}{X_1 + X_2}, \frac{X_1 + X_2}{X_1 + X_2 + X_3}, \dots, \frac{X_1 + \dots + X_{n-1}}{X_1 + \dots + X_n}, X_1 + \dots + X_n。$$

雖上述結果是在假設 $\alpha_i, i = 1, \dots, n$, 為已知之下證出, 但顯然在上述推導中, $\alpha_i, i = 1, \dots, n$, 是否已知對推導過程並未有任何影響, 故此結果對所有 $\alpha_i, i = 1, \dots, n$, 及 β 成立。

我們看到適當地利用巴蘇定理, 可得此一通常以變數代換來證明的結果(見第三章定理3.1及例7.7)。附帶一提, $X_1/(X_1 + X_2)$ 有 $Be(\alpha_1, \alpha_2)$ 分佈, $(X_1 + X_2)/(X_1 + X_2 + X_3)$ 有 $Be(\alpha_1 + \alpha_2, \alpha_3)$ 分佈, 餘類推。在上述證明中, 因 Y_1 與 T_2 獨立, 故我們亦得 $X_1/(X_1 + X_2)$ 與 $X_1 + X_2$ 獨立, 此亦為第三章定理3.1所給之統計上一著名的結果。

底下我們再給一定理3.2不適用之例。

例3.12 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, \theta^2)$ 分佈所產生之隨機樣本, 其中 $\theta > 0$ 。即 X_i 's 之共同 p.d.f. 為 $(\theta\sqrt{2\pi})^{-1} \exp\{x/\theta - x^2/(2\theta^2) - 1/2\}$ 。則 $T = T(\mathbf{X}) = (T_1, T_2) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 為 θ 之一最小充分統計量。但與(2.7)式對照, $C = (w_1(\theta), w_2(\theta)) = (1/\theta, -1/(2\theta^2))$, $\theta > 0$, 其圖形為二維平面上一條曲線, 並未包含任何二維開矩形, 故定理3.2不適用。但 T 究竟是否為完備充分統計量呢?

令

$$g(T) = 2T_1^2 - (n+1)T_2 = 2\left(\sum_{i=1}^n X_i\right)^2 - (n+1)\sum_{i=1}^n X_i^2,$$

則 $E_\theta(g(T)) = 0, \forall \theta > 0$ (習題第8題)。因此 $T(\mathbf{X})$ 並無完備性。讀者是否能看出為何取這樣的 g ?

定理3.1也可用來證明一統計量並非完備。

例3.13 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(\theta - 1/2, \theta + 1/2)$ 分佈所產生之隨機樣本, $-\infty < \theta < \infty$ 。則 $T = T(\mathbf{X}) = (X_{(1)}, X_{(n)})$ 為 θ 之一最小充分統計量。但因 T 與輔助統計量 $X_{(n)} - X_{(1)}$ 並不獨立, 故由巴蘇定理知, T 非完備的。

事實上, 若取 $g(T) = X_{(n)} - X_{(1)} - (n-1)/(n+1)$, 則得

$$\begin{aligned} E_\theta(g(T)) &= \left(\theta - \frac{1}{2} + \frac{n}{n+1}\right) - \left(\theta - \frac{1}{2} + \frac{1}{n}\right) - \frac{n-1}{n+1} \\ &= 0, \quad \theta \in R. \end{aligned}$$

結束本節前，我們請問各位：既然有充分統計量，那是否有“必要統計量”(necessary statistic)呢？事實上是有的。根據Dynkin(1951)，我們有下述定義：一統計量若能寫成每一充分統計量之函數，便稱為必要統計量。比較此定義與最小充分統計量的定義，立即得到下述有趣的定理：一統計量為最小充分統計量，若且唯若它為充分且必要統計量。

習 題

第5.2節

1. 設函數 f 與 g 有相同定義域，且 g 為嚴格單調。試證 f 為 g 之函數。
2. 令 $f(x) = x^3 + 3x^2 + 1$, $g(x) = x^4 + 3x^2 + 1$, $x \geq 0$ 。試證 f 為 g 之函數，且 g 亦為 f 之函數。
3. 令 $f(x) = x^3 + 3x^2 + 1$, $g(z) = x^4 - 2x^2 + 1$, $x \geq 0$ 。試證 g 為 f 之函數，而 f 並非 g 之函數。
4. 試判斷下述各小題中，二統計量是否等價，其中 $n \geq 2$ 。
 - (i) $\prod_{i=1}^n X_i$, $\sum_{i=1}^n \log X_i$, X_i 取值皆為正;
 - (ii) $\sum_{i=1}^n X_i$, $\sum_{i=1}^n \log X_i$, X_i 取值皆為正;
 - (iii) $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^3)$, $(\sum_{i=1}^n X_i, \sum_{i=1}^n (X_i - \bar{X}_n)^3)$ 。

(解. (i) 是, (ii) 是, (iii) 否)
5. 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta)$ 所產生之隨機樣本。試證順序統計量 $(X_{(1)}, \dots, X_{(n)})$ 為 θ 之一充分統計量。
6. 設 $X_1, \dots, X_n$ 為一組隨機樣本, $n \geq 2$ 。試證統計量
 - (i) $(\bar{X}_n, S_n^2)$ 與 $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 等價;
 - (ii) $(X_{(1)}, X_{(n)})$ 與 $(X_{(n)} - X_{(1)}, (X_{(n)} + X_{(1)})/2)$ 等價。
7. 設 X_1, X_2 為由 $\mathcal{P}(\lambda)$ 分佈所產生之隨機樣本, $\lambda > 0$ 。試求在給定 $X_1 + X_2 = t$ 之下, X_1, X_2 之條件分佈，並據此判定 $T(\mathbf{X}) = X_1 + X_2$ 是否為 λ 之一充分統計量。

8. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。試分別以下述三法, 判定 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 為 θ 之一充分統計量。

(i) 試證給定 $T = t$, $X_1, \dots, X_n$ 之條件 p.d.f. 為

$$f_{X_1, \dots, X_n | T}(x_1, \dots, x_n | T = t) = \frac{(n-1)!}{t^{n-1}}, \sum_{i=1}^n x_i = t;$$

(ii) 定理 2.2;

(iii) 定理 2.3。

9. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生隨機樣本, $\theta > 0$ 。試分別以下述二法, 判定 $T(\mathbf{X}) = X_{(n)}$ 是否為 θ 之一充分統計量。

(i) 求在給定 $X_{(n)} = t$ 之下, $X_1, \dots, X_n$ 之條件分佈;

(ii) 利用定理 2.2。

10. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{D}\text{-}\mathcal{U}(1, \theta)$ 分佈所產生之隨機樣本, 其中 θ 為未知且取正整數值。試證 $X_{(n)}$ 為 θ 之一充分統計量。

11. 設 $X_1, \dots, X_n$ 為一組由 $\Gamma(\alpha, \beta)$ 分佈所產生之隨機樣本。試證

(i) 若 α, β 皆未知, 則 $(\prod_{i=1}^n X_i, \sum_{i=1}^n X_i)$ 為 (α, β) 之一充分統計量;

(ii) 若 β 已知, 則 $\prod_{i=1}^n X_i$ 為 α 之一充分統計量;

(iii) 若 α 已知, 則 $\sum_{i=1}^n X_i$ 為 β 之一充分統計量。

12. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{B}e(\alpha, \beta)$ 分佈所產生之隨機樣本。試證

(i) 若 α, β 皆未知, 則 $(\prod_{i=1}^n X_i, \prod_{i=1}^n (1 - X_i))$ 為 (α, β) 之一充分統計量;

(ii) 若 β 已知, 則 $\prod_{i=1}^n X_i$ 為 α 之一充分統計量;

(iii) 若 α 已知, 則 $\prod_{i=1}^n (1 - X_i)$ 為 β 之一充分統計量;

(iv) 若 $\beta = \alpha$, 則 $\prod_{i=1}^n (X_i(1 - X_i))$ 為 α 之一充分統計量。

13. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\alpha, p)$ 所產生之隨機樣本, 其中

$$f(x|\alpha, p) = (1-p)p^{x-\alpha}, x = \alpha, \alpha+1, \dots,$$

$\alpha \in R, 0 < p < 1$ 。試證

(i) 若 α, p 皆未知, 則 $(X_{(1)}, \sum_{i=1}^n X_i)$ 為 (α, p) 之一充分統計量;

(ii) 若 p 已知, 則 $X_{(1)}$ 為 α 之一充分統計量;

(iii) 若 α 已知, 則 $\sum_{i=1}^n X_i$ 為 p 之一充分統計量。

14. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{W}(\alpha, \beta)$ 分佈所產生之隨機樣本。

(i) 設 β 已知, 試求 α 之一充分統計量;

(ii) 設 α 已知, 試問是否存在 β 之一維的充分統計量。

(解. (i) $\prod_{i=1}^n X_i^\beta$, (ii) 只有 $n = 1$ 才存在)

15. 設 $\mathbf{X} = (X_1, \dots, X_n)$ 有 n 維常態分佈。 $E(X_1) = n\mu$, $E(X_i) = 0$, $i = 2, 3, \dots, n$ 。又共變異數矩陣為

$$\Sigma = \begin{pmatrix} n-1+\sigma^2 & -1 & -1 & \cdots & -1 \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

(i) 試求 $\mathbf{X}$ 之聯合 p.d.f.;

(ii) 試證 $\bar{X}_n$ 為 (μ, σ^2) 之一充分統計量。

16. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{P}(\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。試求 θ 之一最小充分統計量。

17. 試證例 2.5 中之 T 為 θ 之一最小充分統計量。並給出所有最小充分統計量。

18. 在例 2.12 中, 試證 $I_{\{-\theta \leq x_{(1)} \leq x_{(n)} \leq \theta\}} = I_{\{\max\{|x_{(1)}|, |x_{(n)}|\} \leq \theta\}}$, 且 $\max\{|X_{(1)}|, |X_{(n)}|\}$ 為 θ 之一最小充分統計量。

19. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta) = \theta^{-2}e^{-(x-\theta)/\theta^2}$, $x \geq \theta$, $\theta > 0$, 所產生之隨機樣本。試求 θ 之一最小充分統計量。

20. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\sigma, \sigma^2)$ 分佈所產生之隨機樣本, $\sigma > 0$ 。試求 σ 之一最小充分統計量。
21. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{C}(\theta, 1)$ 分佈所產生之隨機樣本, $\theta \in R$ 。試求 θ 之一最小充分統計量。
22. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta)$ 所產生之隨機樣本, 其中

$$f(x|\theta) = \begin{cases} \frac{1}{\sigma} e^{-(x-\mu)/\sigma} & , x \geq \mu, \\ 0 & , \text{其他}, \end{cases}$$

$$\theta = (\mu, \sigma), \mu \in R, \sigma > 0。$$

- (i) 若 μ, σ 皆未知, 試求 θ 之一最小充分統計量;
- (ii) 若 σ 已知, 試求 μ 之一最小充分統計量;
- (iii) 若 μ 已知, 試求 σ 之一最小充分統計量。
23. 分別重做第 11、12 及 13 題, 但將充分統計量改為最小充分統計量。
24. 設 $X_1, \dots, X_n, n \geq 2$, 為一組由 p.d.f. $f(x|\theta) = (x/\theta)e^{-x^2/(2\theta)}, x > 0, \theta > 0$, 所產生之隨機樣本。
- (i) 試證 $T(\mathbf{X}) = \sum_{i=1}^n X_i^2$ 為 θ 之一最小充分統計量;
- (ii) 試證 $\bar{X}_n = \sum_{i=1}^n X_i/n$ 並非 θ 之一充分統計量。
25. 設 $X_1, \dots, X_n$ 為獨立的隨機變數, 且 X_k 有波松分佈, 參數為 $(\alpha/k) \cdot (n/(n+\beta))^k, 1 \leq k \leq n$ 。試求 (α, β) 之一最小充分統計量。
26. 試證 (2.5) 式成立。
27. 設隨機變數 X 取值在 $\{1, 2, 3, 4\}$, 其 p.d.f. 為

$$f(x) = \begin{cases} 1/4 & , x = 1, 4, \\ (3-\theta)/12 & , x = 2, \\ (3+\theta)/12 & , x = 3, \end{cases}$$

其中 $\theta \in [0, 1]$ 。令 $T(1) = T(4) = 3, T(2) = 2, T(3) = 4$ 。試以分解定理證明 T 為 θ 之一充分統計量。

28. 試利用分解定理證明定理2.5。
29. 試證定理2.6。
30. 試證定理2.7。
31. 設 $X_1, \dots, X_n$ 為 i.i.d. 的隨機變數, 且 $P(X_1 = \theta) = 1 - P(X_1 = -\theta) = 2/3, \theta \in R \setminus \{0\}$ 。令 M 表 $\{X_1, \dots, X_n\}$ 中正的個數, $Y = |X_1|$ 。
- (i) 試證 M, Y 分別為 $\text{sgn}\theta$ (表 θ 之符號) 及 $|\theta|$ 之充分統計量;
- (ii) 試求 M, Y 之聯合分佈。
32. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta)$ 所產生之隨機樣本, $\theta = (\theta_1, \theta_2)$, 令 $\mathbf{X} = (X_1, \dots, X_n)$ 。設當 θ_2 已知時, $T_1(\mathbf{X})$ 為 θ_1 之一充分統計量, 當 θ_1 已知時, $T_2(\mathbf{X})$ 為 θ_2 之一充分統計量。又設 θ_1, θ_2 獨立地變動 (即取值範圍互不影響), 且 $\{x|f(x|\theta) > 0\}$ 與 θ 無關。
- (i) 試證若 T_1, T_2 皆與 θ_1, θ_2 無關, 則 $(T_1(\mathbf{X}), T_2(\mathbf{X}))$ 為 θ 之一充分統計量;
- (ii) 試給一 $(T_1(\mathbf{X}), T_2(\mathbf{X}))$ 為 θ 之充分統計量, 且當 θ_2 已知時, $T_1(\mathbf{X})$ 為 θ_1 之一充分統計量, 但當 θ_1 已知時, $T_2(\mathbf{X})$ 不為 θ_2 之充分統計量的例子。

第5.3節

1. 設 X_1, X_2 為由 p.d.f. $f(x|\alpha) = \alpha x^{\alpha-1} e^{-x^\alpha}, x > 0, \alpha > 0$, 所產生之隨機樣本。試證 $\log X_1 / \log X_2$ 為一輔助統計量。
2. 設 $X_1, \dots, X_n$ 為一組由一位置族分佈所產生之隨機樣本。令 M 表樣本中位數。試證 $M - \overline{X}_n$ 為一輔助統計量。
3. 設 $X_1, \dots, X_n$ 為一組由位置族分佈所產生之隨機樣本。試證對任意 $\sum_{i=1}^n a_i = 0$ 且 $\sum_{i=1}^n a_i^2 \neq 0, \sum_{i=1}^n a_i X_i$ 為一輔助統計量。
4. 設 $X_1, \dots, X_n$ 為一組由尺度族分佈所產生之隨機樣本。試證只要 $\sum_{i=1}^n a_i^2$ 及 $\sum_{i=1}^n b_i^2$ 皆不等於0, 則 $\sum_{i=1}^n a_i X_i / \sum_{i=1}^n b_i X_i$ 為一輔助統計量。

5. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。先利用巴蘇定理, 證明 X_n 與 $X_{(1)}/X_{(n)}$ 獨立, 再利用此結果求 $E(X_{(1)} \cdot X_{(n)}^{-1})$ 。
6. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta) = (1/\sigma)e^{-(x-\theta)/\sigma}I_{\{x \geq \theta\}}$ 所產生之隨機樣本, $\theta \in R$, 而 $\sigma > 0$ 為一常數。
 - (i) 試證 $X_{(1)}$ 是 θ 之一最小充分統計量, 並求 $X_{(1)}$ 之分佈;
 - (ii) 試證 $X_{(1)}$ 為 θ 之一完備統計量;
 - (iii) 試證 $\sum_{i=1}^n (X_i - X_{(1)})$ 與 $X_{(1)}$ 獨立。
7. 設 N 為一取正整數值之隨機變數, $P(N = i) = p_i, i \geq 1, \sum_{i=1}^{\infty} p_i = 1$ 。又設在給定 $N = n$ 之下, X 有 $\mathcal{B}(n, \theta)$ 分佈。試證 (X, N) 為 θ 之一最小充分統計量, 且 N 為 θ 之一輔助統計量。
8. 試證例 3.12 中, 對 $\forall \theta > 0, E_{\theta}(g(T)) = 0$ 。
9. 試證例 2.18 中之最小充分統計量並無完備性。
10. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\theta, a\theta^2)$ 分佈所產生之隨機樣本, 其中 a 為一已知常數, $\theta > 0$ 。試證 $T = (\bar{X}_n, S_n^2)$ 為 θ 之一充分統計量, 但並無完備性。
11. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{G}e(\theta)$ 分佈所產生之隨機樣本, $0 < \theta < 1$, 令 $\mathbf{X} = (X_1, \dots, X_n)$ 。試證 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 為 θ 之一充分統計量。又試判定 T 是否有完備性。
12. 設 X, Y 皆為取 0, 1, 2 三個值之隨機變數, $P(X = 0) = p, P(X = 1) = 3p, 0 < p < 1/4, P(Y = 0) = p, P(Y = 1) = p^2, 0 < p < 1/2$ 。試分別判定 X, Y 之完備性。
13. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta)$ 所產生之隨機樣本。分別針對下述各 f , 找一完備充分統計量, 或證明這種統計量不存在。
 - (i) $f(x|\theta) = 2x/\theta^2, 0 < x < \theta, \theta > 0$;
 - (ii) $f(x|\theta) = \theta/(1+x)^{1+\theta}, 0 < x < \infty, \theta > 0$;

$$(iii) f(x|\theta) = \theta^x \log \theta / (\theta - 1), 0 < x < 1, \theta > 1;$$

$$(iv) f(x|\theta) = e^{-(x-\theta)} \exp(-e^{-(x-\theta)}), -\infty < x < \infty, -\infty < \theta < \infty;$$

$$(v) f(x|\theta) = \binom{2}{x} \theta^x (1-\theta)^{2-x}, x = 0, 1, 2, 0 \leq \theta \leq 1。$$

14. 設 X 有 $\mathcal{U}[\theta, \theta + 1]$ 分佈, $\theta \in R$ 。試問 X 是否為完備充分統計量。
15. 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, $\mu \in R$, $\sigma > 0$, 令 $\mathbf{X} = (X_1, \dots, X_n)$ 。試證 $n \geq 2$ 時, $\mathbf{X}$ 並非完備統計量。
16. 設 X 之p.d.f.為

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, x = -1, 0, 1, 0 \leq \theta \leq 1。$$

試問

- (i) X 是否為完備充分統計量?
- (ii) $|X|$ 是否為完備充分統計量?
- (iii) $f(x|\theta)$ 是否屬於指數族?
17. 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$, 所產生之隨機樣本。
- (i) 試問 $\sum_{i=1}^n X_i$ 是否為 θ 之一充分統計量?
- (ii) 試求 θ 之一完備充分統計量。
18. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(\theta, 2\theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。試求 θ 之一最小充分統計量, 並問此統計量是否具有完備性。
19. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{P}(\theta)$ 分佈所產生之隨機樣本, $\theta = 1, 2$ 。試證此分佈族並無完備性(本題可與例3.6比較)。
20. 設 $X_1, \dots, X_n$ 為一組由p.d.f. $f(x|\theta) = e^{-(x-\theta)}$ 所產生之隨機樣本, $-\infty < \theta < x < \infty$ 。
- (i) 試證 $X_{(1)}$ 為一最小充分統計量;
- (ii) 利用巴蘇定理證明 $X_{(1)}$ 與 S_n^2 獨立。

21. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x|\theta) = c(\theta)e^{-(x-\theta)^4}$, $-\infty < x < \infty$, $-\infty < \theta < \infty$, 所產生之隨機樣本, 其中 $c(\theta)$ 由 $\int_{-\infty}^{\infty} f(x|\theta)dx = 1$ 決定。令 $T_k = \sum_{i=1}^n X_i^k$, $k = 1, 2, 3$ 。試證
- (i) (T_1, T_2, T_3) 為 θ 之一最小充分統計量;
- (ii) $A_1 = nT_2 - T_1^2$ 及 $A_2 = n^2T_3 - 3nT_2T_1 + 2T_1^3$ 皆為輔助統計量。
22. 設有一骰子, 各面出現的機率分別為 $(1-\theta)/6$, $(1-2\theta)/6$, $(1-3\theta)/6$, $(1+\theta)/6$, $(1+2\theta)/6$, $(1+3\theta)/6$, 其中 $-1/3 < \theta < 1/3$ 。令 $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5, X_6)$ 表投擲 n 次後, 各面出現的次數。令 $Y_1 = X_1 + X_4$, $Y_2 = X_2 + X_5$, $Y_3 = X_3 + X_6$ 。
- (i) 試決定 $\mathbf{X}$ 之分佈;
- (ii) 試決定 (Y_1, Y_2, Y_3) 之分佈, 並問此是否為一輔助統計量。

參考文獻

1. Casella, G. and Berger, R.L.(1990). *Statistical Inference*. Wadsworth and Brooks/Cole Publishing Company, Pacific Grove, California.
2. Dynkin, E.B.(1951). Necessary and sufficient statistics for a family of probability distributions. English translation in *Selected Translations in Mathematical Statistics and Probability*, 1(1961), 23-41。
3. Lehmann, E.L.(1983). *Theory of Point Estimation*. Wadsworth and Brooks/Cole Publishing Company, Pacific Grove, California.
4. Levy, M.S.(1985). A note on nonunique mle's and sufficient statistics. *The American Statistician*, 39, No.1,66.

5. Lindgren, B.W.(1976). *Statistical Theory*, 3rd ed. Macmillan College Publishing Company, New York.
6. Moore D.S.(1971). Maximum likelihood and sufficient statistics. *The American Mathematical Monthly*, 78, 50-52.
7. Sampson, A. and Spencer B.(1976). Sufficiency, minimal sufficiency and the lack thereof. *The American Statistician*, 30, No.1, 34-35.