第一章　充分統計量

㈠基本定義

定義A.1

設$(X_1,\cdots,X_n)$為一隨機向量，函數$T(X_1,\cdots,X_n)$為其一函數，且其不包含任意未知參數，則稱$T(X_1,\cdots,X_n)$為一統計量(Statistic)。

定義A.2

若一統計量在給定統計量的值下，$X_1,\cdots,X_n$的分配與母體參數無關，則稱此統計量為充分統計量(Sufficient Statistic)。

【Remark】：

由〔定義A.2〕可知，若$X_1,\cdots,X_n$為離散型隨機向量，

則若滿足

$P\{X_1=x_1,\cdots,X_n=x_n|T(X_1,\cdots,X_n)=t\}$與未知參數$\underline{\theta}$無關，則其為充分統計量。

若$X_1,\cdots,X_n$為連續型隨機向量，

則$\because$ $P\{T(X_1,X_2,\cdots,X_n)=t\}=0$，則必將〔定義A.2〕的準則予以修正，以達尋求充分統計量的目的。

例題A.1

若$Y_1,Y_2,Y_3 \xrightarrow{i.i.d} Ber(\theta)$，且定義$S=Y_1+Y_2+Y_3$，$T=Y_1Y_2+Y_3$則$S$是否是$\theta$的充分統計量？$T$是否是$\theta$的充分統計量？

📖先建立

$$P\{(y_1,y_2,y_3)|S=s\}=\frac{P\{(y_1,y_2,y_3)\cap S=s\}}{P\{S=s\}}$$

$$=\begin{cases}\dfrac{P\{(y_1,y_2,y_3)\}}{P\{S=s\}}, \text{if } \ y_1+y_2+y_3=s \\ 0 \qquad\qquad , \ \text{其他}\end{cases}$$

如下表：

	機率	S	T	$P\{(y_1,y_2,y_3)\|S=0\}$	$P\{(y_1,y_2,y_3)\|S=1\}$
(0,0,0)	$(1-\theta)^3$	0	0	1	0
(1,0,0)	$\theta(1-\theta)^2$	1	0	0	1/3
(0,1,0)	$\theta(1-\theta)^2$	1	0	0	1/3
(0,0,1)	$\theta(1-\theta)^2$	1	1	0	1/3
(0,1,1)	$\theta^2(1-\theta)$	2	1	0	0
(1,0,1)	$\theta^2(1-\theta)$	2	1	0	0
(1,1,0)	$\theta^2(1-\theta)$	2	1	0	0
(1,1,1)	θ^3	3	2	0	0

	機率	$P\{(y_1,y_2,y_3)\|S=2\}$	$P\{(y_1,y_2,y_3)\|S=3\}$
(0,0,0)	$(1-\theta)^3$	0	0
(1,0,0)	$\theta(1-\theta)^2$	0	0
(0,1,0)	$\theta(1-\theta)^2$	0	0
(0,0,1)	$\theta(1-\theta)^2$	0	0
(0,1,1)	$\theta^2(1-\theta)$	1/3	0
(1,0,1)	$\theta^2(1-\theta)$	1/3	0
(1,1,0)	$\theta^2(1-\theta)$	1/3	0
(1,1,1)	θ^3	0	1

由〔**定義A.2**〕可知，$S=Y_1+Y_2+Y_3$爲參數θ的充分統計量。

又$P\{(y_1=0,y_2=1,y_3=0)|T=0\}$

$$=\frac{P\{(y_1=0,y_2=1,y_3=0)\cap T=0\}}{P\{T=0\}}=\frac{P\{(y_1=0,y_2=1,y_3=0)\}}{P\{T=0\}}$$

$$=\frac{\theta(1-\theta)^2}{(1-\theta)^3+2\theta(1-\theta)^2}=\frac{\theta}{1+\theta}$$

與參數θ有關，$\therefore T=Y_1Y_2+Y_3$不爲參數θ的充分統計量。

例題A.2

若$Y_1,\cdots,Y_n \xrightarrow{i.i.d} \text{Ber}(p)$，

試證明$S=\sum_{i=1}^{n}Y_i$爲θ的充分統計量。

📖 $P\{(Y_1=y_1,\cdots,Y_n=y_n)|S=s\}=\dfrac{P\{(Y_1=y_1,\cdots,Y_n=y_n)\cap S=s\}}{P\{S=s\}}$

$$=\frac{P\{(Y_1=y_1,\cdots,Y_n=y_n)\cap S=s\}}{P\{S=s\}}$$

$$=\begin{cases}\dfrac{1}{\binom{n}{s}}, \text{if } y_1+\cdots+y_n=s\\ 0\ ,\ \text{其他}\end{cases}$$

$\therefore$由〔定義A.2〕可知，$S=\sum_{i=1}^{n}Y_i$爲參數θ的充分統計量。

例題A.3

Let $X_1,X_2 \xrightarrow{i.i.d} f(x)=\begin{cases}e^{-\theta} & ;\ \text{if } x=0\\ \theta e^{-\theta} & ;\ \text{if } x=1\\ 1-(\theta+1)e^{-\theta} & ;\ \text{if } x=2\end{cases}$, show that

X_1+X_2 is not a sufficient statistic for θ. (中山應數所)

$\because P\{(X_1=2,X_2=0)\,|\,X_1+X_2=2\}=\dfrac{P\{X_1=2,X_2=0\}}{P\{X_1+X_2=2\}}$

$$=\frac{\{1-(\theta+1)e^{-\theta}\}e^{-\theta}}{2\left\{\{1-(\theta+1)e^{-\theta}\}e^{-\theta}\right\}+\theta^2e^{-2\theta}}$$

$\therefore X_1+X_2$ 並非 θ 的充分統計量。

例題A.4

Let X and Y be random variables having joint density $f(x,y)$ below.

$f(x,y)$	$y=-1$	$y=1$
$x=-1$	$1-3\theta+\theta^2$	$\theta-\theta^2$
$x=1$	$\theta-\theta^2$	$\theta+\theta^2$

Determine θ so that X and Y are independent.
Show that $X+Y$ is a sufficient statistic for θ. (逢甲統精所)

$\because P\{X=1,Y=1\}=\theta+\theta^2$

又 $P\{X=1\}=2\theta=P\{Y=1\}$

則若 $X \coprod Y \Leftrightarrow P\{X=1,Y=1\}=P\{X=1\}P\{Y=1\}(2\times 2\text{ table})$

$$\Leftrightarrow \theta+\theta^2=4\theta^2 \Leftrightarrow \theta=0,\frac{1}{3}$$

$$\because P\{(X=x,Y=y)\,|\,X+Y=-2\}=\begin{cases}1 & ,\text{ if } (x,y)=(-1,-1)\\ 0 & ,\text{ 其他}\end{cases}$$

$$P\{(X=x,Y=y)\,|\,X+Y=0\}=\begin{cases}\frac{1}{2} & ,\text{ if } (x,y)=(-1,1),(1,-1)\\ 0 & ,\text{ 其他}\end{cases}$$

$$P\{(X=x,Y=y)\,|\,X+Y=2\}=\begin{cases}1 & ,\text{ if } (x,y)=(1,1)\\ 0 & ,\text{ 其他}\end{cases}$$

則根據〔**定義A.2**〕可知，$X+Y$ 為 θ 的充分統計量。

定理A.3 Neyman-Fisher分解定理 (Neyman-Fisher factorization theorem)

若 $X_1, X_2, \cdots, X_n \xrightarrow{i.i.d} f(x;\theta)$；

且若存在一統計量 $T(X_1, X_2, \cdots, X_n)$，使得

$$f(x_1, x_2, \cdots x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = h(x_1, x_2, \cdots, x_n) g(T(x_1, x_2, \cdots, x_n);\theta)$$

$\Leftrightarrow$ 統計量 $T(X_1, X_2, \cdots, X_n)$ 為參數 θ 的充分統計量。

【證明】：

僅證明 $X_1, \cdots, X_n$ 為離散型隨機向量情況。

$(\Leftarrow)$

$T(X_1, X_2, \cdots, X_n)$ 為參數 θ 的充分統計量。

則 $P\{X_1 = x_1, \cdots, X_n = x_n | T(X_1, \cdots, X_n) = t\} = h(X_1, \cdots, X_n)$

與未知參數 θ 無關。

$$\therefore \frac{P\{X_1 = x_1, \cdots, X_n = x_n \cap T(X_1, \cdots, X_n) = t\}}{P\{T(X_1, \cdots X_n) = t\}}$$

$$= \begin{cases} \dfrac{P\{X_1 = x_1, \cdots, X_n = x_n\}}{P\{T(X_1, \cdots X_n) = t\}}, & \text{if } T(X_1, \cdots, X_n) = t \\ 0, & \text{if } T(X_1, \cdots, X_n) \neq t \end{cases}$$

$$= h(x_1, \cdots, x_n)$$

令 $g(T(x_1, \cdots, x_n);\theta) = P\{T(X_1, \cdots, X_n) = t\}$

則 $P\{X_1 = x_1, \cdots, X_n = x_n\} = P\{T(X_1, \cdots X_n) = t\} h(x_1, \cdots, x_n)$

$$= g(T(x_1, \cdots, x_n);\theta) h(x_1, \cdots, x_n)$$

$(\Rightarrow)$

令 $P\{X_1 = x_1, \cdots, X_n = x_n\} = f(x_1, x_2, \cdots x_n;\ \theta)$

$= h(x_1, x_2, \cdots, x_n) g(T(x_1, x_2, \cdots, x_n);\theta)$ 成立

則 $P\{T(X_1, \cdots X_n) = t\}$

$$= \sum_{\{T(x_1, \cdots, x_n) = t\}} h(x_1, \cdots, x_n) g(T(x_1, \cdots, x_n);\theta)$$

$$= g(t;\theta) \sum_{\{T(x_1, \cdots, x_n) = t\}} h(x_1, \cdots, x_n)$$

$$\therefore P\{X_1=x_1,\cdots,X_n=x_n \mid T(X_1,\cdots,X_n)=t\}$$
$$=\frac{P\{X_1=x_1,\cdots,X_n=x_n,T(X_1,\cdots,X_n)=t\}}{P\{T(X_1,\cdots X_n)=t\}}$$
$$=\begin{cases}\dfrac{P\{X_1=x_1,\cdots,X_n=x_n\}}{P\{T(X_1,\cdots X_n)=t\}}, & \text{if } T(X_1,\cdots,X_n)=t\\ 0, & \text{if } T(X_1,\cdots,X_n)\neq t\end{cases}$$
$$=\begin{cases}\dfrac{h(x_1,\cdots,x_n)}{\sum_{\{T(x_1,\cdots,x_n)=t\}}h(x_1,\cdots,x_n)}, & \text{if } T(X_1,\cdots,X_n)=t\\ 0, & \text{if } T(X_1,\cdots,X_n)\neq t\end{cases}$$

例題A.5

若(1) $X_1,\cdots,X_n \xrightarrow{i.i.d}$ Poisson (θ)，

(2) $X_1,\cdots,X_n \xrightarrow{i.i.d}$ Ber (θ)，

試分別利用分解定理求 θ 的充分統計量。

📖 (1) $\because X_1,\cdots,X_n \xrightarrow{i.i.d} \text{Poisson}(\theta)$

$$\therefore f(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}=\left(\frac{1}{\prod_{i=1}^{n}x_i!}\right)\left(e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}\right)$$

可令 $h(x_1,\cdots,x_n)=\dfrac{1}{\prod_{i=1}^{n}x_i!}, g\left(\sum_{i=1}^{n}x_i;\theta\right)=e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}$

則根據〔Neyman-Fisher分解定理〕可知，$\sum_{i=1}^{n}X_i$ 爲 θ 的充分統計量。

(2) $\because X_1,\cdots,X_n \xrightarrow{i.i.d} \text{Ber}(\theta)$

$$\therefore f(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\theta^{\sum_{i=1}^{n}x_i}(1-\theta)^{n-\sum_{i=1}^{n}x_i}=\left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n}x_i}(1-\theta)^n$$

可令 $h(x_1,\cdots,x_n)=1, g\left(\sum_{i=1}^{n}x_i;\theta\right)=\left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n}x_i}(1-\theta)^n$

則根據〔Neyman-Fisher分解定理〕可知，$\sum_{i=1}^{n}X_i$ 爲 θ 的充分統計量。

定義A.4　指標函數(indicator function)

若一函數 $I_A(x)$ 滿足 $I_A(x)=\begin{cases}1, & \text{if } x\in A\\ 0, & \text{otherwise}\end{cases}$

則稱函數 $I_A(x)$ 爲指標函數，事件 A 爲指標事件。

【Remark】：

指標函數的特性

(1) $E(I_A(x))=P(A)$, $Var(I_A(x))=P(A)(1-P(A))$

(2)若A與B爲兩指標事件，則 $I_A(x)I_B(x)=I_{A\cap B}(x)$

【證明】：

(1)參見〔機率論重點整理〕〔第5章〕

(2) $I_A(x)\ I_B(x)=\begin{cases}1, \text{if } x\in A \text{ and } x\in B\\ 0, \text{otherwise}\end{cases}=I_{A\cap B}(x)$

若於 X 的p.d.f 中，x_i 的範圍與 θ 有關，則指標函數的建立有利於充分性的探討。

定理A.5

若 $X_1,\cdots,X_n \xrightarrow{i.i.d} f(x;\underline{\theta})$；其中 $\underline{\theta}=(\theta_1,\cdots,\theta_r)^T$

若 $f(x_1,\cdots x_n;\underline{\theta})=\prod_{i=1}^{n}f(x_i;\underline{\theta})=g(T_1(\underline{x}),\cdots,T_k(\underline{x});\underline{\theta})h(\underline{x})$

則 $(T_1(\underline{X}),\cdots,T_k(\underline{X}))$ 是 $\underline{\theta}$ 的(結合)充分統計量(joint sufficient statistic)；

其中 $\underline{X}=(X_1,\cdots,X_n)$

【Remark】：

由〔定理A.5〕可知，$\left(X_1,\cdots,X_n\right)$為$\underline{\theta}$的(結合)充分統計量。

【說例】：

$\because X_1,\cdots,X_n \xrightarrow{i.i.d} \text{Poisson}(\theta)$，

$$\therefore f(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}=\left(\frac{1}{\prod_{i=1}^{n} x_i!}\right)\left(e^{-n\theta}\theta^{x_1}\theta^{x_2}\cdots\theta^{x_n}\right)$$

可令$h\left(x_1,\cdots,x_n\right)=\dfrac{1}{\prod_{i=1}^{n} x_i!}$，　$T_1(\underline{x})=x_1, T_2(\underline{x})=x_2,\cdots T_n(\underline{x})=x_n$

則根據〔定理A.5〕可知，$\left(X_1,\cdots,X_n\right)$為$\underline{\theta}$的(結合)充分統計量。

且$\underline{T}(\underline{X})=(T_1(\underline{X}),\cdots,T_k(\underline{X}))$為$\underline{\theta}$的(結合)充分統計量。

定理A.6

於〔定理A.5〕中，若存在一向量值函數ϕ，滿足

$$\underline{T}^*(\underline{X})=\phi(\underline{T}(\underline{X}))=(T_1^*(\underline{X}),\cdots,T_m^*(\underline{X}))$$

且ϕ為1對1函數，則$\underline{T}^*(\underline{X})$亦為$\underline{\theta}$的(結合)充分統計量。

【證明】：

$\because \underline{T}(\underline{X})=(T_1(\underline{X}),\cdots,T_k(\underline{X}))$為$\underline{\theta}$的(結合)充分統計量。

$$\begin{aligned}\therefore f(x_1,\cdots,x_n;\theta)&=\prod_{i=1}^{n} f(x_i;\theta)=g(T_1(\underline{x}),\cdots,T_k(\underline{x});\underline{\theta})\,h(\underline{x})\\&=g(\underline{T}(\underline{x});\underline{\theta})\,h(\underline{x})\end{aligned}$$

$\because \underline{T}^*(\underline{X})=\phi(\underline{T}(\underline{X}))$為1對1函數，則存在$\phi^{-1}$使得

$\underline{T}(\underline{X})=\phi^{-1}(\underline{T}^*(\underline{X}))$

$$\begin{aligned}\text{則 } f(x_1,\cdots,x_n;\theta)&=g(\underline{T}(\underline{x});\underline{\theta})\,h(\underline{x})=g(\phi^{-1}(\underline{T}^*(\underline{X}));\underline{\theta})\,h(\underline{x})\\&\equiv g^*(T^*(\underline{x});\underline{\theta})\,h(\underline{x})\end{aligned}$$

則由〔定理A.5〕可知，$\underline{T}^*(\underline{X})$亦為$\underline{\theta}$的(結合)充分統計量。

【Remark】：

〔定理A.6〕中，函數ϕ條件未必要1對1，其證明超過本書水準。

例題A.6

(1) Let $X_1,\cdots,X_n \xrightarrow{i.i.d} U(0\,;\theta)$,$\theta>0$ find a sufficient statistic for θ.

(2) Let $X_1,\cdots,X_n \xrightarrow{i.i.d} U(\theta_1\,;\theta_2)$, $\theta_2>\theta_1$, find a sufficient statistic for $\underline{\theta}=(\theta_1,\theta_2)$.

(1) $\because X_1,\cdots,X_n \xrightarrow{i.i.d} f(x\,;\theta)=\begin{cases}\dfrac{1}{\theta}, \text{ if } 0<x<\theta\\ 0\ , \text{otherwise}\end{cases}$

又於$f(x\,;\theta)$中，x的範圍與θ有關。

則將$f(x\,;\theta)$改寫為$f(x\,;\theta)=\dfrac{1}{\theta}I_{(0,\theta)}(x)$

$$\therefore f(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}I_{(0,\theta)}(x_i)=\frac{1}{\theta^n}\prod_{i=1}^{n}I_{(0,\theta)}(x_i)$$

$$\because \prod_{i=1}^{n}I_{(0,\theta)}(x_i)=\begin{cases}1, (0<x_1<\theta)\cap(0<x_2<\theta)\cap\cdots\cap(0<x_n<\theta)\\ 0, \text{otherwise}\end{cases}$$

$$=\begin{cases}1,\ \text{if }\ 0<x_{(1)}\le x_{(2)}\cdots\le x_{(n)}<\theta\\ 0\ ,\ \text{otherwise}\end{cases}=I_{(x_{(n)},\infty)}(\theta)$$

其中$X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$為$(X_1,\cdots,X_n)$的秩序統計量

則$f(x_1,\cdots,x_n;\theta)=\dfrac{1}{\theta^n}\displaystyle\prod_{i=1}^{n}I_{(0,\theta)}(x_i)=\dfrac{1}{\theta^n}I_{(x_{(n)},\infty)}(\theta)$

令$h(x_1,\cdots,x_n)=1, g(x_{(n)},\theta)=\dfrac{1}{\theta^n}I_{(x_{(n)},\infty)}(\theta)$

則根據Neyman-Fisher分解定理可知，$X_{(n)}=\max\limits_i X_i$為$\theta$的充分統計量。

(2) $\because X_1,\cdots,X_n \xrightarrow{i.i.d} f(x;\theta_1,\theta_2)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & \text{if } \theta_1<x<\theta_2\\ 0, & \text{otherwise}\end{cases}$

又於 $f(x;\theta_1,\theta_2)$ 中，x的範圍與 θ_1,θ_2 有關。

則將 $f(x;\theta_1,\theta_2)$ 改寫爲 $f(x;\theta_1,\theta_2)=\dfrac{1}{\theta_2-\theta_1}I_{(\theta_1,\theta_2)}(x)$

$$\therefore f(x_1,\cdots,x_n;\theta_1,\theta_2)=\prod_{i=1}^{n}\frac{1}{\theta_2-\theta_1}I_{(\theta_1,\theta_2)}(x_i)=\frac{1}{(\theta_2-\theta_1)^n}\prod_{i=1}^{n}I_{(\theta_1,\theta_2)}(x_i)$$

$$\because \prod_{i=1}^{n}I_{(\theta_1,\theta_2)}(x_i)$$

$$=\begin{cases}1, & (\theta_1<x_1<\theta_2)\cap(\theta_1<x_2<\theta_2)\cap\cdots\cap(\theta_1<x_n<\theta_2)\\ 0, & \text{otherwise}\end{cases}$$

$$=\begin{cases}1, & \text{if } \theta_1<x_{(1)}\le x_{(2)}\cdots\le x_{(n)}<\theta_2\\ 0, & \text{otherwise}\end{cases}$$

$$=I_{(-\infty,x_{(1)})}(\theta_1)I_{(x_{(n)},\infty)}(\theta_2)$$

$$\text{則 } f(x_1,\cdots,x_n;\theta_1,\theta_2)=\prod_{i=1}^{n}\frac{1}{\theta_2-\theta_1}I_{(\theta_1,\theta_2)}(x_i)$$

$$=\frac{1}{(\theta_2-\theta_1)^n}I_{(-\infty,x_{(1)})}(\theta_1)I_{(x_{(n)},\infty)}(\theta_2)$$

令 $h(x_1,\cdots,x_n)=1, g(x_{(1)},x_{(n)};\theta_1,\theta_2)=\dfrac{1}{(\theta_2-\theta_1)^n}I_{(-\infty,x_{(1)})}(\theta_1)I_{(x_{(n)},\infty)}(\theta_2)$

則根據〔**定理A.5**〕可知，$(X_{(1)},X_{(n)})$ 爲 (θ_1,θ_2) 的(結合)充分統計量。

例題A.7

Let $X_1,\cdots,X_n \xrightarrow{i.i.d} N(\theta_1;\theta_2)$, $\theta_1\in\Re$, $\theta_2>0$, find a sufficient statistic for $\underline{\theta}=(\theta_1,\theta_2)$.

📖 $\because X_1,\cdots,X_n \xrightarrow{i.i.d} f(x;\theta_1,\theta_2)=\dfrac{1}{\sqrt{2\pi\theta_2}}\exp\left\{-\dfrac{(x-\theta_1)^2}{2\theta_2}\right\}, -\infty<x<\infty$

$$\therefore f(x_1,\cdots,x_n;\theta_1,\theta_2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta_2}}\exp\left\{-\frac{(x_i-\theta_1)^2}{2\theta_2}\right\}$$

$$=\frac{1}{(2\pi\theta_2)^{\frac{n}{2}}}\exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\theta_1)^2}{2\theta_2}\right\}$$

$$=\frac{1}{(2\pi\theta_2)^{\frac{n}{2}}}\exp\left\{-\frac{\sum_{i=1}^{n}x_i^2}{2\theta_2}+\frac{\theta_1\sum_{i=1}^{n}x_i}{\theta_2}-\frac{n\theta_1^2}{2\theta_2}\right\}$$

令 $h(\underset{\sim}{x})=1, g\left(\sum_{i=1}^{n}x_i,\sum_{i=1}^{n}x_i^2,\theta_1,\theta_2\right)$

$$=\frac{1}{(2\pi\theta_2)^{\frac{n}{2}}}\exp\left\{-\frac{\sum_{i=1}^{n}x_i^2}{2\theta_2}+\frac{\theta_1\sum_{i=1}^{n}x_i}{\theta_2}-\frac{n\theta_1^2}{2\theta_2}\right\}$$

則根據〔**定理A.5**〕可知，$\left(\sum_{i=1}^{n}x_i^2,\sum_{i=1}^{n}x_i\right)$ 爲 (θ_1,θ_2) 的(結合)充分統計量。

【Remark】：

於〔**例題A.7**〕中，$\because \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，$s^2=\frac{1}{n-1}\left\{\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right\}$

則可知 $\left(\sum_{i=1}^{n}x_i^2,\sum_{i=1}^{n}x_i\right)\to(\bar{x},s^2)$ 爲1對1轉換，則 $(\bar{X},S^2)$ 亦爲 (θ_1,θ_2) 的(結合)充分統計量。

例題A.8

Let $X_1,\cdots,X_n$ be i.i.d with p.d.f

$$f(x;\theta)=\begin{cases}\dfrac{1}{2}\theta e^{-\theta x}, 0<x<\infty \\ 1 \quad , -\dfrac{1}{2}\le x\le 0 \ , \ \theta>0 \\ 0 \quad , x<\dfrac{-1}{2}\end{cases}$$

Find a sufficient statistic for θ. (中央統研所)

$$\because f(x_1,\cdots,x_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\left(\frac{1}{2}\theta\right)^{\sum_{i=1}^{n}I_{(0,\infty)}(x_i)}\exp\left\{-\theta\sum_{i=1}^{n}x_iI_{(0,\infty)}(x_i)\right\}$$

$$=\left(\frac{1}{2}\right)^{\sum_{i=1}^{n}I_{(0,\infty)}(x_i)}\left\{\left\{\theta^{\sum_{i=1}^{n}I_{(0,\infty)}(x_i)}\right\}\exp\left\{-\theta\sum_{i=1}^{n}x_iI_{(0,\infty)}(x_i)\right\}\right\}$$

則可令 $h(\underline{x})=\left(\frac{1}{2}\right)^{\sum_{i=1}^{n}I_{(0,\infty)}(x_i)}$

$$g\left(\sum_{i=1}^{n}I_{(0,\infty)}(x_i),\sum_{i=1}^{n}x_iI_{(0,\infty)}(x_i);\theta\right)$$

$$=\left\{\left\{\theta^{\sum_{i=1}^{n}I_{(0,\infty)}(x_i)}\right\}\exp\left\{-\theta\sum_{i=1}^{n}x_iI_{(0,\infty)}(x_i)\right\}\right\}$$

則根據〔定理A.5〕可知，$\left(\sum_{i=1}^{n}I_{(0,\infty)}(x_i),\sum_{i=1}^{n}x_iI_{(0,\infty)}(x_i)\right)$為$\theta$的(結合)充分統計量。

㈡完備性(Completeness)與相關議題

定義B.1

若隨機變數X，其p.d.f爲 $f(x;\theta)$ ，且其參數空間爲Θ ；則可定義集合$\left\{f(x;\theta)\middle|\theta\in\Theta\right\}$爲隨機變數$X$的**機率密度函數族(family of p.d.f)**。

定義B.2

若隨機變數X，其p.d,f族爲$\{f(x;\theta)|\theta\in\Theta\}$ ，若於此族中，滿足

$$E_\theta\{g(X)\}=0,\ \forall\theta\in\Theta\Rightarrow g(X)\equiv 0$$

$$(\text{或}P_\theta\{g(X)=0\}=1,\forall\theta\in\Theta)$$

則稱此族是**完備的(completeness)**。

【Remark】：

於點估計的理論中，完備性的定義相當於是在此族中，

0的唯一的不偏估計量爲0。

完備性爲一p.d.f族的特性。

定義B.3

若一統計量$T(\underline{X})$ ，其p.d.f爲 $f(t;\theta)$ ，若其對應的p.d.f族爲$\{f(t;\theta)|\theta\in\Theta\}$ 爲完備的，則稱統計量$T(\underline{X})$ 爲**完備統計量(complete statistic)**。

例題B.1

(1)Let $X_1,\cdots,X_n \xrightarrow{i.i.d}$ Ber(θ), $0<\theta<1$, to show that $\sum_{i=1}^{n}X_i$ is complete.