ch6：数据简化原理

6.2 充分性原理

6.2 充分性原理

参数 θ 的一个**充分统计量**在某种意义上提炼了样本中有关 θ 的全部信息，即除充分统计量的值以外，样本中其余信息不能再提供关于 θ 的任何信息. 这就是数据简化充分性原理的基本思想.

充分性原理：如果 $T(\boldsymbol{X})$ 是 θ 的一个充分统计量，则 θ 的任意依赖于样本 $\boldsymbol{X}$ 的推断都可以经由值 $T(\boldsymbol{X})$ 完成，即，如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是满足 $T(\boldsymbol{x})=T(\boldsymbol{y})$ 的两个样本点，则不论观测到的是 $\boldsymbol{X}=\boldsymbol{x}$ 还是 $\boldsymbol{X}=\boldsymbol{y}$，关于 θ 的推断都完全相同.

本节我们主要考察充分统计量与充分性原理.

6.2.1 充分统计量

定义6.2.1

充分统计量的严格定义如下：

定义 6.2.1 如果样本 $\boldsymbol{X}$ 在已知统计量 $T(\boldsymbol{X})$ 取值时的条件分布与 θ 无关，则称统计量 $T(\boldsymbol{X})$ 是 θ 的**充分统计量**（sufficient statistic).

注意:

- 这个想法是，如果我知道$T(X)$，则关于θ的推论仅取决于$T(X)$的值。 因此，给定$T(X)$，我可以丢弃实际数据。

定理6.2.2

> **定理 6.2.2** 设 $p(\boldsymbol{x}|\theta)$ 为样本 $\boldsymbol{X}$ 的联合概率密度（或质量）函数，$q(t|\theta)$ 为 $T(\boldsymbol{X})$ 的概率密度（或质量）函数. 如果对样本空间中的任意 $\boldsymbol{x}$，比值 $p(\boldsymbol{x}|\theta)/q(T(\boldsymbol{x})|\theta)$ 都是 θ 的常函数，则 $T(\boldsymbol{X})$ 是 θ 的充分统计量.

例6.2.3

例 6.2.3（二项充分统计量） 设 $X_1,\cdots,X_n$ 是参数为 θ，$0<\theta<1$ 的 Bernoulli 随机样本，我们将证明 $T(\boldsymbol{X})=X_1+\cdots+X_n$ 是 θ 的充分统计量. 注意 $T(\boldsymbol{X})$ 记录了取值为 1 的 X_i 的数目，所以 $T(\boldsymbol{X})$ 服从参数为 (n,θ) 的二项分布，从而概率质量函数之比为

$$\frac{p(\boldsymbol{x}|\theta)}{q(T(\boldsymbol{x})|\theta)}=\frac{\prod\theta^{x_i}(1-\theta)^{1-x_i}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}}\qquad(令\ t=\sum x_i)$$

$$=\frac{\theta^{\sum x_i}(1-\theta)^{\sum(1-x_i)}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} \qquad (\text{因为}\prod\theta^{x_i}=\theta^{\sum x_i})$$

$$=\frac{\theta^t(1-\theta)^{n-t}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}}$$

$$=\frac{1}{\binom{n}{t}}$$

$$=\frac{1}{\binom{n}{\sum x_i}}$$

显然与 θ 无关，故由定理 6.2.2 知 $T(\boldsymbol{X})$是 θ 的充分统计量．其意义如下：Bernoulli 样本中 1 的个数概括了样本中有关 θ 的全部信息，样本的其他特征，比如 X_3 的取值等都不能提供关于 θ 的更多信息． ‖

注意

根据充分统计量的定义直接求充分统计量并不普遍．因为为了运用定义，我们必须首先猜测充分统计量 $T(\boldsymbol{X})$，求出 $T(\boldsymbol{X})$的概率密度（或质量）函数，然后验证概率密度（或质量）函数之比与 θ 无关．其中第一步需要凭借很好的直觉，第二步有时又需进行冗长的分析计算．令人庆幸的是，借助下列由 Halmos and Savage（1949）给出的定理，我们只需简单考察样本 pdf 或 pmf 便可求得充分统计量[⊖].

定理6.2.6（因子分解定理）

定理 6.2.6（因子分解定理） 设 $f(\boldsymbol{x}\mid\theta)$ 为样本 $\boldsymbol{X}$ 的联合概率密度（或质量）函数，统计量 $T(\boldsymbol{X})$是 θ 的充分统计量当且仅当存在函数 $g(t\mid\theta)$和 $h(\boldsymbol{x})$，使得对任意样本点 $\boldsymbol{x}$ 以及参数 θ，都有

(6.2.3) $$f(\boldsymbol{x}\mid\theta)=g(T(\boldsymbol{x})\mid\theta)h(\boldsymbol{x}).$$

证明 我们只证明离散分布的情形.

假设 $T(\boldsymbol{X})$是 θ 的充分统计量．令 $g(t\mid\theta)=P_\theta(T(\boldsymbol{X})=t), h(\boldsymbol{x})=P(\boldsymbol{X}=\boldsymbol{x}\mid T(\boldsymbol{X})=T(\boldsymbol{x}))$．由于 $T(\boldsymbol{X})$是充分统计量，用于定义 $h(\boldsymbol{x})$的条件概率不依赖于 θ，因此 $h(\boldsymbol{x})$和 $g(t\mid\theta)$的定义合法且满足：

$$f(\boldsymbol{x}\mid\theta)=P_\theta(\boldsymbol{X}=\boldsymbol{x})$$

$$
\begin{aligned}
&= P_\theta(\boldsymbol{X}=\boldsymbol{x} \text{ 且 } T(\boldsymbol{X})=T(\boldsymbol{x})) \\
&= P_\theta(T(\boldsymbol{X})=T(\boldsymbol{x}))P(\boldsymbol{X}=\boldsymbol{x} \mid T(\boldsymbol{X})=T(\boldsymbol{x})) \quad (\text{根据 } T(\boldsymbol{X}) \text{ 的充分性}) \\
&= g(T(\boldsymbol{x})\mid\theta)h(\boldsymbol{x})
\end{aligned}
$$

这就得到式（6.2.3）．由上式最后两行亦知：

$$P_\theta(T(\boldsymbol{X})=T(\boldsymbol{x}))=g(T(\boldsymbol{x})\mid\theta)$$

故 $g(T(\boldsymbol{x})\mid\theta)$是 $T(\boldsymbol{X})$的概率质量函数.

现在假设分解式（6.2.3）成立．令 $q(t\mid\theta)$为 $T(\boldsymbol{X})$的概率质量函数，为证明 $T(\boldsymbol{X})$是 θ 的充分统计量我们只需考察比值 $f(\boldsymbol{x})/q(T(\boldsymbol{x})\mid\theta)$．定义 $A_{T(\boldsymbol{x})}=\{\boldsymbol{y}: T(\boldsymbol{y})=T(\boldsymbol{x})\}$，则

$$
\begin{aligned}
\frac{f(\boldsymbol{x}\mid\theta)}{q(T(\boldsymbol{x})\mid\theta)} &= \frac{g(T(\boldsymbol{x})\mid\theta)h(\boldsymbol{x})}{q(T(\boldsymbol{x})\mid\theta)} && (\text{根据式}(6.2.3)) \\
&= \frac{g(T(\boldsymbol{x})\mid\theta)h(\boldsymbol{x})}{\sum A_{T(\boldsymbol{x})}g(T(\boldsymbol{y})\mid\theta)h(\boldsymbol{y})} && (\text{根据 } T \text{ 的概率质量函数的定义}) \\
&= \frac{g(T(\boldsymbol{x})\mid\theta)h(\boldsymbol{x})}{g(T(\boldsymbol{x})\mid\theta)\sum A_{T(\boldsymbol{x})}h(\boldsymbol{y})} && (\text{因为 } T \text{ 在 } A_{T(\boldsymbol{x})} \text{ 上的取值恒定}) \\
&= \frac{h(\boldsymbol{x})}{\sum A_{T(\boldsymbol{x})}h(\boldsymbol{y})}
\end{aligned}
$$

显然该比值与 θ 无关，故由定理 6.2.2 可知 $T(\boldsymbol{X})$是 θ 的充分统计量. □

例子

- 令$X_1,...,X_n$服从独立同分布的伯努利分布$Bernoulli(\theta)$，证明$\sum x_i$是θ充分统计量

例子

- 令$X_1,...,X_n$服从独立同分布的泊松分布$Possion(\theta)$，证明$\sum x_i$和$\bar{x}$是θ充分统计量

例子

- 令$X_1,...,X_n$服从独立同分布的正态分布$N(\mu,\sigma^2)$，σ^2已知，证明$\bar{x}$是μ充分统计量

例子

令$X_1,...,X_n$服从独立同分布的正态分布$N(\mu,\sigma^2)$，σ^2未知知，证明$(\bar{x},S^2)$是(μ,σ^2)充分统计量

例子（均匀充分统计量）

- 令$X_1,...,X_n$取自在$1,...,\theta$上离散均匀分布的总体，证明$T(X)=\max X_i$是θ的充分统计量

定理6.2.10

定理 6.2.10 设随机样本 $X_1, \cdots, X_n$ 取自概率密度（或质量）函数为 $f(x|\theta)$ 的总体，其中 $f(x|\theta)$ 属指数族概率密度（或质量）函数，其定义为：

$$f(x \mid \boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right)$$

其中 $\theta=(\theta_1,\theta_2,\cdots,\theta_d), d\leqslant k$. 则

$$T(\boldsymbol{X}) = \left(\sum_{j=1}^{n} t_1(X_j),\cdots,\sum_{j=1}^{n} t_k(X_j)\right)$$

是 $\boldsymbol{\theta}$ 的充分统计量.

例子

Example

Let $X_1,\ldots,X_n$ be i.i.d. $\mathrm{N}(\theta, a\theta)$, where $a>0$ is known. (The variance is proportional to the mean.)

$$f(x_i|\theta) = (2\pi a\theta)^{-1/2}\exp\left[-(x_i-\theta)^2/2a\theta\right]$$

$$=(2\pi a\theta)^{-1/2}\exp\left[-\left(x_i^2-2x_i\theta+\theta^2\right)/2a\theta\right]$$

$$=(2\pi a\theta)^{-1/2}\exp\left[-\frac{x_i^2}{2a\theta}+\frac{x_i}{a}-\frac{\theta}{2a}\right].$$

Set

$h(x_i)=\exp(x_i/a)$,

$c(\theta)=(2\pi a\theta)^{-1/2}\exp(-\theta/2a)$,

$w_1(\theta)=-1/(2a\theta)$ and $t_1(x_i)=x_i^2$. So $T(\mathbf{x})=\sum x_i^2$ is sufficient for θ.

注意

- 充分统计量的任何一对一函数都是充分统计量

6.2.2 极小充分统计量

定义6.2.11

定义 6.2.11 称充分统计量 $T(\boldsymbol{X})$ 是**极小充分统计量**（minimal sufficient statistic），如果对其余任一充分统计量 $T'(\boldsymbol{X})$，$T(\boldsymbol{X})$ 都是 $T'(\boldsymbol{X})$ 的函数.

例6.2.12

例 6.2.12（两个正态充分统计量） 例 6.2.4 所考察的概率模型包含 $\mathrm{n}(\mu,\sigma^2)$ 随机样本 X_1，…，X_n，其中参数 σ^2 已知. 由分解式（6.2.4）可知 $T(\boldsymbol{X})=\overline{X}$ 是 μ 的充分统计量. 此外分解式（6.2.5）亦成立（只不过现在已知 σ^2），因此 $T'(\boldsymbol{X})=(\overline{X},S^2)$ 也是 μ 的充分统计量. 显然 $T(\boldsymbol{X})$ 所进行的数据简化优于 $T'(\boldsymbol{X})$，因为单凭 $T(\boldsymbol{X})$ 我们并不知道样本方差 S^2. 事实上我们可以利用函数 $r(a,b)=a$ 将 $T(\boldsymbol{x})$ 写成 $T'(\boldsymbol{x})$ 的函数：$T(\boldsymbol{x})=\bar{x}=r(\bar{x},s^2)=r(T'(\boldsymbol{x}))$. 由于 $T(\boldsymbol{X})$ 和 $T'(\boldsymbol{X})$ 均为 μ 的充分统计量，它们都包含同样多的关于 μ 的信息. 又因为已知 σ^2，故样本方差 S^2 的取值这个额外信息并不能为我们带来更多关于 μ 的信息. 当然，如果 σ^2 未知，则如例 6.2.9 所示，$T(\boldsymbol{X})=\overline{X}$ 不再是充分统计量，且此时 $T'(\boldsymbol{X})$ 所包含的关于（μ，σ^2）的信息要多于 $T(\boldsymbol{X})$. ‖

定理6.2.13——求极小充分统计量

定理 6.2.13 设 $f(\boldsymbol{x}|\theta)$ 是样本 $\boldsymbol{X}$ 的概率密度（或质量）函数. 如果存在函数 $T(\boldsymbol{x})$ 使得对任意两个样本点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，比值 $f(\boldsymbol{x}|\theta)/f(\boldsymbol{y}|\theta)$ 是 θ 的常函数当且仅当 $T(\boldsymbol{x})=T(\boldsymbol{y})$，则 $T(\boldsymbol{X})$ 是 θ 的极小充分统计量.

证明 为简化证明，不妨设对任意 $\boldsymbol{x}\in\mathcal{X}$ 以及 θ 都有 $f(\boldsymbol{x}|\theta)>0$.

先证 $T(\boldsymbol{x})$ 是充分统计量. 令 $\mathcal{T}=\{t:$ 存在 $\boldsymbol{x}\in\mathcal{X}$ 使得 $t=T(\boldsymbol{x})\}$ 为 $\mathcal{X}$ 在 $T(\boldsymbol{x})$ 下的象，由 $T(\boldsymbol{x})$ 确定的分划集记作 $A_t=\{\boldsymbol{x}:T(\boldsymbol{x})=t\}$，$t\in\mathcal{T}$. 对每个 A_t，选定某一 $\boldsymbol{x}_t\in A_t$. 则对任意 $\boldsymbol{x}\in\mathcal{X}$，$\boldsymbol{x}_{T(\boldsymbol{x})}$ 都与 $\boldsymbol{x}$ 落在同一分划集（比如 A_t）当中，即 $T(\boldsymbol{x})=T(\boldsymbol{x}_{T(\boldsymbol{x})})$，于是 $f(\boldsymbol{x}|\theta)/f(\boldsymbol{x}_{T(\boldsymbol{x})}|\theta)$ 是 θ 的常函数. 因此，若令 $\mathcal{X}$ 上的函数 $h(\boldsymbol{x})=f(\boldsymbol{x}|\theta)/f(\boldsymbol{x}_{T(\boldsymbol{x})}|\theta)$，则 h 与 θ 无关. 若再令 $\mathcal{T}$ 上的函数 $g(t|\theta)=f(\boldsymbol{x}_t|\theta)$，则有

$$f(\boldsymbol{x}|\theta)=\frac{f(\boldsymbol{x}_{T(\boldsymbol{x})}|\theta)f(\boldsymbol{x}|\theta)}{f(\boldsymbol{x}_{T(\boldsymbol{x})}|\theta)}=g(T(\boldsymbol{x})|\theta)h(\boldsymbol{x})$$

根据因子分解定理，$T(\boldsymbol{X})$ 是 θ 的充分统计量.

下证 $T(\boldsymbol{X})$ 是极小充分统计量. 设 $T'(\boldsymbol{X})$ 是任一充分统计量. 根据因子分解定理，存在函数 g' 和 h' 使得 $f(\boldsymbol{x}|\theta)=g'(T'(\boldsymbol{x})|\theta)h'(\boldsymbol{x})$. 设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是满足 $T'(\boldsymbol{x})=T'(\boldsymbol{y})$ 的两个样本点，则

$$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)}=\frac{g'(T'(\boldsymbol{x})|\theta)h'(\boldsymbol{x})}{g'(T'(\boldsymbol{y})|\theta)h'(\boldsymbol{y})}=\frac{h'(\boldsymbol{x})}{h'(\boldsymbol{y})}$$

显然与 θ 无关，所以由定理的假设可知 $T(\boldsymbol{x})=T(\boldsymbol{y})$. 故 $T(\boldsymbol{x})$ 是 $T'(\boldsymbol{x})$ 的函数，且 $T(\boldsymbol{x})$ 是极小充分统计量. □

例6.2.14

例 6.2.14（正态极小充分统计量） 设 X_1，…，X_n 是 $\mathrm{n}(\mu,\sigma^2)$ 随机样本，其中 μ 和 σ^2 均未知，$(\bar{x},s_x^2)$ 和 $(\bar{y},s_y^2)$ 分别为样本点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的样本均值与样本方差. 则由式（6.2.5）可知，概率密度函数之比为

$$\begin{aligned}\frac{f(\boldsymbol{x}|\mu,\sigma^2)}{f(\boldsymbol{y}|\mu,\sigma^2)}&=\frac{(2\pi\sigma^2)^{-n/2}\exp(-[n(\bar{x}-\mu)^2+(n-1)s_x^2]/(2\sigma^2))}{(2\pi\sigma^2)^{-n/2}\exp(-[n(\bar{y}-\mu)^2+(n-1)s_y^2]/(2\sigma^2))}\\&=\exp(-[n(\bar{x}^2-\bar{y}^2)+2n\mu(\bar{x}-\bar{y})-(n-1)(s_x^2-s_y^2)]/(2\sigma^2))\end{aligned}$$

显然，该比值是 μ 和 σ^2 的常函数当且仅当 $\bar{x}=\bar{y}$ 且 $s_x^2=s_y^2$. 于是根据定理 6.2.13，（$\overline{X}$，S^2）是（μ，σ^2）的极小充分统计量. ‖

6.2.3 辅助统计量

定义6.2.16

定义 6.2.16 如果统计量 $S(\boldsymbol{X})$ 的分布与 θ 无关，则称 $S(\boldsymbol{X})$ 为**辅助统计量（ancillary statistic）**.

单个的辅助统计量显然不包含任何关于 θ 的信息，它所观测的随机变量服从已知的某种概率分布且与 θ 毫无关系. 然而，辅助统计量一旦与其他统计量联合起来就有可能包含有关 θ 的信息，我们将在下一节讨论这种现象. 现在我们看几个辅助统计量的例子.

定义6.2.17

例 6.2.17（均匀辅助统计量） 同例 6.2.15，设随机样本 $X_1, \cdots, X_n$ 取自在 $(\theta,\theta+1), -\infty<\theta<\infty$ 区间上均匀分布的总体，$X_{(1)}<\cdots<X_{(n)}$ 为样本的次序统计量. 下面我们证明极差 $R=X_{(n)}-X_{(1)}$ 是辅助统计量，这只需说明 R 的概率密度函数与 θ 无关即可.

回忆每个 X_i 的累积分布函数为

$$F(x|\theta)=\begin{cases}0 & x\leqslant\theta\\ 2-\theta & \theta<x<\theta+1\\ 1 & \theta+1\leqslant x\end{cases}$$

于是 $X_{(1)}$ 和 $X_{(n)}$ 的联合概率密度函数如式（5.5.7）所示，为：

$$g(x_{(1)},x_{(n)}|\theta)=\begin{cases}n(n-1)(x_{(n)}-x_{(1)})^{n-2} & \theta<x_{(1)}<x_{(n)}<\theta+1\\ 0 & \text{否则}\end{cases}$$

做变量替换 $R=X_{(n)}-X_{(1)}$ 以及 $M=(X_{(1)}+X_{(n)})/2$，其逆变换为 $X_{(1)}=(2M-R)/2$，

$X_{(n)}=(2M+R)/2$，Jacobi 行列式等于 1. 所以 R 和 M 的联合概率密度函数为

$$h(r,m|\theta)=\begin{cases}n(n-1)r^{n-2} & 0<r<1,\ \theta+(r/2)<m<\theta+1-(r/2)\\ 0 & \text{否则}\end{cases}$$

（注意 $h(r,m|\theta)$ 取正值的范围）于是，R 的概率密度函数为

$$\begin{aligned}h(r\mid\theta)&=\int_{\theta+(r/2)}^{\theta+1-(r/2)} n(n-1)r^{n-2}\,\mathrm{d}m\\&=n(n-1)r^{n-2}(1-r),0<r<1\end{aligned}$$

显然是参数为 $\alpha=n-1$，$\beta=2$ 的贝塔概率密度函数. 注意该概率密度函数与 θ 无关，即 R 的分布与 θ 无关，故 R 是辅助统计量. ‖

例6.2.18

例 6.2.18（位置族辅助统计量） 设随机样本 $X_1, \cdots, X_n$ 取自累积分布函数为 $F(x-\theta)$，$-\infty<\theta<\infty$ 的位置参数族总体，下证极差 $R=X_{(n)}-X_{(1)}$ 是辅助统计量. 根据定理 3.5.6，若令 $X_1=Z_1+\theta, \cdots, X_n=Z_n+\theta$，则 $Z_1, \cdots, Z_n$ 是取自 $F(x)$（即取 $\theta=0$）总体的随机样本. 于是，极差统计量 R 的累积分布函数为

$$\begin{aligned} F_R(r|\theta) &= P_\theta(R\leqslant r) \\ &= P_\theta(\max_i X_i - \min_i X_i \leqslant r) \\ &= P_\theta(\max_i (Z_i+\theta) - \min_i (Z_i+\theta) \leqslant r) \\ &= P_\theta(\max_i Z_i - \min_i Z_i + \theta - \theta \leqslant r) \\ &= P_\theta(\max_i Z_i - \min_i Z_i \leqslant r) \end{aligned}$$

由于 $Z_1, \cdots, Z_n$ 不依赖于 θ，所以上式最后得到的概率与 θ 无关，即 R 的累积分布函数与 θ 无关. 故 R 是辅助统计量. ||

例6.2.19

例 6.2.19（尺度族辅助统计量） 尺度参数族也有一类辅助统计量. 设随机样本 $X_1, \cdots, X_n$ 取自累积分布函数为 $F(x/\sigma)$，$\sigma>0$ 的尺度参数族总体，则所有只通过 $X_1/X_n, \cdots, X_{n-1}/X_n$ 这 $n-1$ 个值与样本关联的统计量都是辅助统计量，例如

$$\frac{X_1+\cdots+X_n}{X_n}=\frac{X_1}{X_n}+\cdots+\frac{X_{n-1}}{X_n}+1$$

事实上，若令 $X_i=\sigma Z_i$，则 $Z_1, \cdots, Z_n$ 是取自 $F(x)$（即取 $\sigma=1$）总体的随机样本. 于是，$X_1/X_n, \cdots, X_{n-1}/X_n$ 的累积分布函数为

$$\begin{aligned} F(y_1,\cdots,y_{n-1}|\sigma) &= P_\sigma(X_1/X_n\leqslant y_1,\cdots,X_{n-1}/X_n\leqslant y_{n-1}) \\ &= P_\sigma(\sigma Z_1/(\sigma Z_n)\leqslant y_1,\cdots,\sigma Z_{n-1}/(\sigma Z_n)\leqslant y_{n-1}) \\ &= P_\sigma(Z_1/Z_n\leqslant y_1,\cdots,Z_{n-1}/Z_n\leqslant y_{n-1}) \end{aligned}$$

由于 $Z_1, \cdots, Z_n$ 不依赖于 σ，所以上式最后得到的概率与 σ 无关，即 $X_1/X_n, \cdots,$ X_{n-1}/X_n 的分布与 σ 无关，其任意函数的分布也均与 σ 无关.

特别地，设 X_1, X_2 是 $\mathrm{n}(0,\sigma^2)$ 随机样本，根据上述讨论可知对任意 σ，X_1/X_2 的分布均相同. 而由例 4.3.6 我们知道当 $\sigma=1$ 时 X_1/X_2 服从参数为（0，1）的 Cauchy 分布，因此对任意 $\sigma>0$，X_1/X_2 都服从 Cauchy 分布. ||

6.2.4 充分统计量、辅助统计量、完全统计量

定义6.2.21

定义 6.2.21 设 $f(t|\theta)$ 是统计量 $T(\boldsymbol{X})$ 的概率密度（或质量）函数，如果满足：对任意 θ 都有 $\mathrm{E}_\theta g(T)=0$，那么对任意 θ 都有 $P_\theta(g(T)=0)=1$，则称该概率分布族是**完全**（**complete**）的，或称 $T(\boldsymbol{X})$ 是**完全统计量**（**complete statistic**）.

注意

注意，完全性是整个概率分布族而非某个特定分布的性质．例如，若 X 服从 $\mathrm{n}(0,1)$ 分布，令 $g(x)=x$，则 $\mathrm{E}g(X)=\mathrm{E}X=0$；但函数 $g(x)=x$ 满足 $P(g(X)=0)=P(X=0)=0$，而非 1．不过这里讨论的是一个特定的分布，而非一族分布．如果 X 服从 $\mathrm{n}(\theta,1)$ 分布，其中 $-\infty<\theta<\infty$，我们将发现 X 的所有函数 $g(x)$ 当中，满足对任意 θ 都有 $\mathrm{E}_\theta g(X)=0$ 的只有一个，该函数对任意 θ 都满足 $P_\theta(g(X)=0)=1$．因此，$\mathrm{n}(\theta,1)(-\infty<\theta<\infty)$ 分布族是完全的．

例6.2.22

例 6.2.22（二项完全充分统计量） 设 T 服从参数为（n，p）的二项分布，其中 $0<p<1$，函数 g 满足 $\mathrm{E}_p g(T)=0$．则对任意 p，$0<p<1$，都有

$$0=\mathrm{E}_p g(T)=\sum_{t=0}^{n} g(t)\binom{n}{t}p^t(1-p)^{n-t}$$
$$=(1-p)^n\sum_{t=0}^{n} g(t)\binom{n}{t}\left(\frac{p}{1-p}\right)^t$$

显然因子 $(1-p)^n$ 恒不为 0，于是对任意 r，$0<r<\infty$，有

$$0=\sum_{t=0}^{n} g(t)\binom{n}{t}\left(\frac{p}{1-p}\right)^t=\sum_{t=0}^{n} g(t)\binom{n}{t}r^t$$

其中等式最右端是 r 的一个 n 次多项式，r^t 项的系数为 $g(t)\binom{n}{t}$．由于该多项式对任意 r 都为 0，故其系数均为 0．而 $\binom{n}{t}$ 项不可能得 0，因此对任意 $t=0$，…，n 都有 $g(t)=0$．又因为 T 的取值只能是 0，1，…，n，所以对任意 p 都有 $P_p(g(T)=0)=1$，故 T 是完全统计量． ||

定理6.2.25

定理 6.2.25（指数族的完全统计量） 设随机变量 X_1，…，X_n 取自概率密度（或质量）函数为

$$f(x\mid\boldsymbol{\theta})=h(x)c(\boldsymbol{\theta})\exp\left(\sum_{j=1}^{k} w(\theta_j)t_j(x)\right)\qquad(6.2.7)$$

的指数族总体，其中 $\boldsymbol{\theta}=(\theta_1,\theta_2,\cdots,\theta_k)$．如果参数空间 Θ 包含 $\mathbf{R}^k$ 的开集，则统计量

$$T(\boldsymbol{X})=\left(\sum_{i=1}^{n} t_1(X_i),\sum_{i=1}^{n} t_2(X_i),\cdots,\sum_{i=1}^{n} t_k(X_i)\right)$$

是完全统计量．

例子

- 令$X_1,...,X_n$独立同分布$N(\mu,\sigma^2)$，证明$(\sum x_i,\sum x_i^2)$是完备充分统计量

例子

- 令$X_1,...,X_n$独立同分布$N(\theta,\theta^2)$，证明$(\sum x_i,\sum x_i^2)$不是完备充分统计量

例子

- 令$X_1,...,X_n$取自在$1,...,\theta$上离散均匀分布的总体，证明$T(X)=\max X_i$是θ的完备充分统计量

一些结论

1

1. If T is complete and $S=\Psi(T)$ is a function of T, then S is also complete.

Proof: If not, then $E_\theta\left[g(S)\right]=0$ for all θ. But then $E_\theta\left[g\left(\Psi(T)\right)\right]=0$ for all θ, and T cannot be complete, a contradiction.

(Consider the composite function $g^*=g(\Psi)$).

2

2. Complete sufficient statistic is also minimal sufficient statistic

3

3. An ancillary statistic cannot be complete.

Proof: If S is ancillary then its distribution does not depend on θ.

So $E_\theta\left[g(S)\right]=E\left[g(S)\right]=C$ is constant in θ. So $g^*(S)=g(S)-C$ has $E\left[g^*(S)\right]=0$ for all θ. Thus S is not complete.

4. If a function of T is ancillary, then T cannot be complete.

Proof: Suppose that $\Psi(T)$ is ancillary, whereas T is complete. Completeness of T implies that $\Psi(T)$ must be complete as well. Since an ancillary statistic cannot be complete, we have a contradiction.

Note: This implies that no complete sufficient statistic exists for the $U(\theta,\theta+1)$ family. This follows since $T(\mathbf{X})=\left(X_{(1)},X_{(n)}-X_{(1)}\right)$ is minimal sufficient, and thus a function of every sufficient statistic, while the function of T, $X_{(n)}-X_{(1)}$, is ancillary.

5. An estimator $g(T)$ for a parameter $\Psi(\theta)$ (based on T only) is called *unbiased* if $E_\theta(g(T)) = \Psi(\theta)$ for all θ.

If T is complete then only one unbiased estimators based on T is possible.

Proof: If there were two unbiased estimator based on T, then $E_\theta(g(T)) = E_\theta(h(T)) = \Psi(\theta)$, but then $E_\theta(g^*(T)) = 0$ for $g^*(t) = g(t) - h(t)$.

In Chapter 7, we will show that this estimator based on a complete sufficient statistic is the best possible unbiased estimator in a certain sense.

定理6.2.24

定理 6.2.24（Basu 定理） 设 $T(\boldsymbol{X})$是完全的极小充分统计量，则 $T(\boldsymbol{X})$与任意辅助统计量都独立.

证明 我们只证明离散分布的情形.

设 $S(\boldsymbol{X})$是任一辅助统计量，则 $P(S(\boldsymbol{X})=s)$与 θ 无关. 又因为 $T(\boldsymbol{X})$是充分统计量（回忆其定义!），故条件概率

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

也与 θ 无关. 因此，为证 $S(\boldsymbol{X})$与 $T(\boldsymbol{X})$独立，只需证明对任意 $t\in\mathcal{T}$都有

$$P(S(\boldsymbol{X})=s|T(\boldsymbol{X})=t)=P(S(\boldsymbol{X})=s) \tag{6.2.6}$$

而

$$P(S(\boldsymbol{X})=s)=\sum_{t\in\mathcal{T}}P(S(\boldsymbol{X})=s\mid T(\boldsymbol{X})=t)P_\theta(T(\boldsymbol{X})=t).$$

又由 $\sum_{t\in\mathcal{T}}P_\theta(T(\boldsymbol{X})=t)=1$，有

$$P(S(\boldsymbol{X})=s)=\sum_{t\in\mathcal{T}}P(S(\boldsymbol{X})=s)P_\theta(T(\boldsymbol{X})=t)$$

所以，若令

$$g(t)=P(S(\boldsymbol{X})=s|T(\boldsymbol{X})=t)-P(S(\boldsymbol{X})=s)$$

则由上面两式可知，对任意 θ 有

$$\mathrm{E}_\theta g(T)=\sum_{t\in\mathcal{T}}g(t)P_\theta(T(\boldsymbol{X})=t)=0$$

因为 $T(\boldsymbol{X})$是完全统计量，故对任意 $t\in\mathcal{T}$都有 $g(t)=0$，这就证得式（6.2.6）. □

例6.2.26

例 6.2.26（Basu 定理的应用-Ⅰ） 设随机样本 $X_1,\cdots,X_n$ 取自参数为 θ 的指数族总体，考虑如何计算

$$g(\boldsymbol{X})=\frac{X_n}{X_1+\cdots+X_n}$$

的期望. 由于指数分布属于尺度参数族，故由例 6.2.19 可知 $g(\boldsymbol{X})$是辅助统计量. 令 $t(x)=x$，则由定理 6.2.25，

$$T(\boldsymbol{X})=\sum_{i=1}^n X_i$$

是 θ 的完全统计量，又由定理 6.2.10 可知 $T(\boldsymbol{X})$是 θ 的充分统计量. 于是，根据 Basu 定理（后面我们将说明无需验证 $T(\boldsymbol{X})$的极小性，尽管根据定理 6.2.13 这是显然的），$T(\boldsymbol{X})$与 $g(\boldsymbol{X})$独立，这样就有

$$\theta=\mathrm{E}_\theta X_n=\mathrm{E}_\theta T(\boldsymbol{X})g(\boldsymbol{X})=(\mathrm{E}_\theta T(\boldsymbol{X}))(\mathrm{E}_\theta g(\boldsymbol{X}))=n\theta\mathrm{E}_\theta g(\boldsymbol{X})$$

故对任意 θ，有 $\mathrm{E}_\theta g(\boldsymbol{X})=n^{-1}$. ‖

6.3似然原理

6.3.1 似然函数

定义6.3.1

定义 6.3.1 设 $f(\boldsymbol{x}|\theta)$为样本 $\boldsymbol{X}=(X_1,\cdots,X_n)$的联合概率密度（或质量）函数，如果观测到 $\boldsymbol{X}=\boldsymbol{x}$，则称 θ 的函数

$$L(\theta|\boldsymbol{x})=f(\boldsymbol{x}|\theta)$$

为似然函数（likelihood function）.

似然原理

似然原理：设样本点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 满足 $L(\theta|\boldsymbol{x})$ 与 $L(\theta|\boldsymbol{y})$ 成比例，即存在某常数 $C(\boldsymbol{x},\boldsymbol{y})$ 使得对任意 θ 有

(6.3.1) $$L(\theta|\boldsymbol{x})=C(\boldsymbol{x},\boldsymbol{y})L(\theta|\boldsymbol{y})$$

则由 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 出发所作的关于 θ 的推断完全相同.

注意，对于不同的样本对 $(\boldsymbol{x},\boldsymbol{y})$，式（6.3.1）中的常数 $C(\boldsymbol{x},\boldsymbol{y})$ 不一定相同，但 $C(\boldsymbol{x},\boldsymbol{y})$ 始终与 θ 无关.

特别地若 $C(\boldsymbol{x},\boldsymbol{y})=1$，则似然原理表明，如果两样本点导出相同的似然函数，则它们所包含的关于 θ 的信息完全相同. 似然原理本质上揭示了似然函数可以用于比较不同参数值的似真程度：如果 $L(\theta_2|\boldsymbol{x})=2L(\theta_1|\boldsymbol{x})$，则从某种意义上说 θ_2 的似真性是 θ_1 的两倍. 如果式（6.3.1）成立，则 $L(\theta_2|\boldsymbol{y})=2L(\theta_1|\boldsymbol{y})$，因此不论观测到的是 $\boldsymbol{x}$ 还是 $\boldsymbol{y}$，我们都可以断定 θ_2 的似真性是 θ_1 的两倍.

我们之所以谨慎地使用“似真性”一词而非“可能性”，是因为考虑到 θ 是一个定值（虽然其值未知）. 此外还需注意，尽管 $f(\boldsymbol{x}|\theta)$ 作为 $\boldsymbol{x}$ 的函数是一个概率密度函数，我们并不能保证 $L(\theta|\boldsymbol{x})$ 作为 θ 的函数也是概率密度函数.

有一种推断的方式叫做**信仰推断**（fiducial inference)，有时将似然函数解释为 θ 的概率，即将 $L(\theta|\boldsymbol{x})$ 乘以 $M(\boldsymbol{x})=\left(\int_{-\infty}^{\infty}L(\theta\mid\boldsymbol{x})\mathrm{d}\theta\right)^{-1}$（若参数空间可数，则应将积分换成求和)，于是 $M(\boldsymbol{x})L(\theta|\boldsymbol{x})$ 就是 θ 的概率密度函数（当然需要假定 $M(\boldsymbol{x})$ 有限!). 显然，若 $L(\theta|\boldsymbol{x})$ 和 $L(\theta|\boldsymbol{y})$ 满足式（6.3.1)，则它们必将导出相同的概率密度函数（常数 $C(\boldsymbol{x},\boldsymbol{y})$ 在标准化时将被消去). 信仰推断理论虽然至今未能得到大部分统计学家的认同，却有着很长的历史，可以追溯到 Fisher（1930）关于**逆概率**（inverse probability)（概率积分变换的应用之一）的工作. 下面我们具体地计算一个信仰分布（fiducial distribution).

6.3.2 形式化的似然原理

形式化的充分性原理　考察试验 $E=(\boldsymbol{X},\theta,\{f(\boldsymbol{x}|\theta)\})$，设 $T(\boldsymbol{X})$ 是 θ 的充分统计量. 如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是满足 $T(\boldsymbol{x})=T(\boldsymbol{y})$ 的样本点，则 $\mathrm{Ev}(E,\boldsymbol{x})=\mathrm{Ev}(E,\boldsymbol{y})$.

6.2 节所给的充分性原理中不涉及试验，与之相比**形式化**的充分性原理更进一步，它表明充分统计量相等时得到的证据亦相等. 似然原理可以由形式化的充分性原理与下列原理推出：

条件原理　设试验 $E_1=(\boldsymbol{X}_1,\theta,\{f_1(\boldsymbol{x}_1|\theta)\})$ 和 $E_2=(\boldsymbol{X}_2,\theta,\{f_2(\boldsymbol{x}_2|\theta)\})$ 有公共的未知参数 θ. 考察混合试验，该试验首先观测到随机变量 J，其中 $P(J=1)=P(J=2)=\frac{1}{2}$（与 θ，X_1 和 X_2 都无关)，然后执行试验 E_J. 混合试验可以形式化地写作 $E^*=(\boldsymbol{X}^*,\theta,\{f^*(\boldsymbol{x}^*|\theta)\})$，其中 $\boldsymbol{X}^*=(j,\boldsymbol{X}_j)$ 且 $f^*(\boldsymbol{x}^*|\theta)=f^*((j,\boldsymbol{x}_j)|\theta)=\frac{1}{2}f_j(\boldsymbol{x}_j|\theta)$. 则

(6.3.2) $$\mathrm{Ev}(E^*,(j,\boldsymbol{x}_j))=\mathrm{Ev}(E_j,\boldsymbol{x}_j)$$

条件原理表明，如果从两试验中随机选取一个执行并观测到样本数据 $\boldsymbol{x}$，则由此得到的有关 θ 的信息**只依赖于被执行的试验**，即与一开始就确定（并非随机选择）执行该试验且观测到样本 $\boldsymbol{x}$ 所能得到的信息相同；这个试验被执行并未增加、减少或者更改我们关于 θ 的信息.

形式化的似然原理　设试验 $E_1=(\boldsymbol{X}_1,\theta,\{f_1(\boldsymbol{x}_1|\theta)\})$ 和 $E_2=(\boldsymbol{X}_2,\theta,\{f_2(\boldsymbol{x}_2|\theta)\})$ 有公共的未知参数 θ. $\boldsymbol{x}_1^*$ 和 $\boldsymbol{x}_2^*$ 分别是 E_1 和 E_2 的样本点，且满足：存在只与

$\boldsymbol{x}_1^*$ 和 $\boldsymbol{x}_2^*$ 有关的常数 C，使得对任意 θ，都有

(6.3.3) $$L(\theta|\boldsymbol{x}_2^*)=CL(\theta|\boldsymbol{x}_1^*)$$

则

$$\mathrm{Ev}(E_1,\boldsymbol{x}_1^*)=\mathrm{Ev}(E_2,\boldsymbol{x}_2^*)$$

形式化的似然原理与 6.3.1 节中给出的似然原理有所不同：似然原理仅考察一个试验，而形式化的似然原理考察两个试验．不过，似然原理可以由形式化的似然原理推出，只需令 E_2 等于 E_1 即可．如此看来，形式化的似然原理中提到两个试验似乎有些不自然．下面我们给出它的一个重要推论，其证明留作练习（习题 6.32）.

似然原理的推论 设 $E=(\boldsymbol{x},\theta,\{f(\boldsymbol{x}|\theta)\})$为一试验，则 $\mathrm{Ev}(E,\boldsymbol{x})$只通过 $L(\theta|\boldsymbol{x})$与 E 和 $\boldsymbol{x}$ 关联.

现在给出 Birnbaum 定理，然后考察它的一些惊人的应用.

定理6.3.6

定理 6.3.6（Birnbaum 定理） 形式化的似然原理可以由形式化的充分性原理以及条件原理推出，反之亦然.

证明 我们只给出证明概要，细节留作习题 6.33. 设 E_1，E_2，$\boldsymbol{x}_1^*$ 和 $\boldsymbol{x}_2^*$ 的定义同形式化的似然原理，E^* 表示条件原理中的混合试验．在 E^* 的样本空间上定义下列统计量

$$T(j,\boldsymbol{X}_j)=\begin{cases}(1,\boldsymbol{x}_1^*) & \text{如果 } j=1 \text{ 且 } \boldsymbol{x}_1=\boldsymbol{x}_1^*\text{，或者 } j=2 \text{ 且 } \boldsymbol{x}_2=\boldsymbol{x}_2^* \\ (j,\boldsymbol{x}_j) & \text{否则}\end{cases}$$

根据因子分解定理可知 $T(J,\boldsymbol{X}_J)$是试验 E^* 中的一个充分统计量．再由形式化的充分性原理，有

(6.3.4) $$\mathrm{Ev}(E^*,(1,\boldsymbol{x}_1^*))=\mathrm{Ev}(E^*,(2,\boldsymbol{x}_2^*))$$

而由条件原理，有

$$\mathrm{Ev}(E^*,(1,\boldsymbol{x}_1^*))=\mathrm{Ev}(E_1,\boldsymbol{x}_1^*)$$

(6.3.5) $$\mathrm{Ev}(E^*,(2,\boldsymbol{x}_2^*))=\mathrm{Ev}(E_2,\boldsymbol{x}_2^*)$$

于是 $\mathrm{Ev}(E_1,\boldsymbol{x}_1^*)=\mathrm{Ev}(E_2,\boldsymbol{x}_2^*)$，这就证得形式化的似然原理.

反过来，设其中一个试验为 E^*，另一个为 E_j．可以证明 $\mathrm{Ev}(E^*,(j,\boldsymbol{x}_j))=\mathrm{Ev}(E_j,\boldsymbol{x}_j)$，即条件原理成立．此外若充分统计量 $T(\boldsymbol{X})$满足 $T(\boldsymbol{x})=T(\boldsymbol{y})$，则似然函数成比例，于是由形式化的似然原理可知 $\mathrm{Ev}(E,\boldsymbol{x})=\mathrm{Ev}(E,\boldsymbol{y})$，即证得形式化的充分性原理. □

6.4 同变性原理

前两节介绍的数据简化原理采用的是下列方式：指定样本函数 $T(\boldsymbol{X})$，如果样本点 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 满足 $T(\boldsymbol{x})=T(\boldsymbol{y})$，则无论观测到的是 $\boldsymbol{x}$ 还是 $\boldsymbol{y}$，我们所推得的关于 θ 的信息都相同. 对于充分性原理，函数 $T(\boldsymbol{x})$ 是充分统计量；对于似然原理，$T(\boldsymbol{x})$ 的"值"就是与 $L(\theta|\boldsymbol{x})$ 成比例的全体似然函数. 而同变性原理所描述的数据简化方式则有所不同：指定函数 $T(\boldsymbol{x})$ 后，如果 $T(\boldsymbol{x})=T(\boldsymbol{y})$，则同变性原理要求观测到 $\boldsymbol{x}$ 时所作的推断与观测到 $\boldsymbol{y}$ 时所作的推断之间存在**某种联系**、但可以不同. 对推断过程的这种限制有时也能帮助我们达到数据简化的目的. ⊖

尽管本节要介绍的数据简化技术统称为同变性原理，它实际上包含了两种不同的同变思想.

第一种同变称作**度量同变**，它要求关于参数 θ 所作的推断不应依赖于所选用的测量尺度. 例如，两名林务员拟估计森林中树木树干的平均直径，第一名林务员以英寸为单位记录树干直径，第二名则以米为单位记录. 现在要求他们同时以英寸为单位估计平均直径（第二名林务员显然很方便以米为单位估计平均直径，然后再将其换算为英寸单位)，度量同变要求两人得到相同的估计值. 毫无疑问，几乎所有人都会认同这种同变是合理的.

第二种同变实际上是不变，可以称作**形式不变**，它要求数学模型形式结构相同的两个推断问题可以运用相同的推断过程. 模型中的元素相同，都包含 Θ——参数空间，$\{f(\boldsymbol{x}|\theta):\theta\in\Theta\}$——样本概率密度（或质量）函数集，以及**正确推断与错误推断集**. 其中最后一项前面未作介绍，本节我们假设可能的推断集就是 Θ，即关于

θ 的一次推断就是从 Θ 中选取一个元素作为 θ 真实值的一个估计. 形式不变性只考察试验的数学结构，而不关心其实际背景. 例如，在不同的试验中 Θ 都可能为 $\{\theta:\theta>0\}$，试验 1 中的 θ 可能表示美国一打鸡蛋的平均价格（单位：美分)，试验 2 中的 θ 可能表示肯尼亚长颈鹿的平均身高（单位：米). 然而，根据形式不变性两试验的参数空间完全相同，因为它们在数学上都表示同一个实数集.

同变性原理　设 $\boldsymbol{T}=g(\boldsymbol{X})$ 是一个度量尺度变换且满足：$\boldsymbol{Y}$ 的模型与 $\boldsymbol{X}$ 的模型具有相同的形式结构，则推断方法应该同时满足度量同变与形式不变.

定义6.4.2

定义 6.4.2　称样本空间 $\mathcal{X}$ 到自身上的一集函数 $\{g(\boldsymbol{x}):g\in\mathcal{G}\}$ 为 $\mathcal{X}$ 的**变换群**（**group of transformations**)，如果

(i)（逆）对任意 $g\in\mathcal{G}$，存在 $g'\in\mathcal{G}$ 使得：对任意 $\boldsymbol{x}\in\mathcal{X}$，都有 $g'(g(\boldsymbol{x}))=\boldsymbol{x}$；

(ii)（复合）对任意 $g\in\mathcal{G}$ 以及 $g'\in\mathcal{G}$，存在 $g''\in\mathcal{G}$ 使得：对任意 $\boldsymbol{x}\in\boldsymbol{X}$，都有 $g'(g(\boldsymbol{x}))=g''(\boldsymbol{x})$；

(iii)（单位元）$\mathcal{G}$ 中存在单位元 $e(\boldsymbol{x})$，其定义为：$e(\boldsymbol{x})=\boldsymbol{x}$.

其中第三条有时作为群定义中的一部分，有时被略去. 它可以由（i)，（ii）推得，

定义6.4.4

定义 6.4.4　设 $\mathcal{F}=\{f(\boldsymbol{x}|\theta):\theta\in\Theta\}$ 是 $\boldsymbol{X}$ 的概率密度（或质量）函数族，$\mathcal{G}$ 是样本空间 $\mathcal{X}$ 的变换群. 如果对任意的 $\theta\in\Theta$ 和 $g\in\mathcal{G}$，都存在唯一的 $\theta'\in\Theta$ 使得：若 $\boldsymbol{X}$ 服从 $f(\boldsymbol{x}|\theta)$ 分布，则 $\boldsymbol{Y}=g(\boldsymbol{X})$ 服从 $f(\boldsymbol{y}|\theta')$ 分布，则称 $\mathcal{F}$ **在群** $\mathcal{G}$ **下不变**（**invariant under the group** $\mathcal{G}$).

6.5 习题

6.3

6.3 设随机样本 X_1，…，X_n 取自概率密度函数为

$$f(x|\mu,\sigma^2)=\frac{1}{\sigma}e^{-(x-\mu)/\sigma},\ \mu<x<+\infty,0<\sigma<+\infty$$

的总体，求(μ,σ^2)的一个二维充分统计量.

6.12

6.12 在许多问题中一个最自然的辅助统计量是**样本的大小**，例如，设 N 是取值为 1，2，…的随机变量，对应的概率分别为 p_1，p_2…，其中 $\sum p_i=1$. 假设观测到 $N=n$，即执行 n 次成功概率为 p 的 Bernoulli 试验，观测到成功试验有 X 次.

(a) 证明：数对(X,N)是 θ 的极小充分统计量，N 是 θ 的辅助统计量（注意此处与 4.4 节层次模型的相似之处）；

(b) 证明 X/N 是 θ 的无偏估计，且其方差为 $\theta(1-\theta)\mathrm{E}(1/N)$.

6.15

6.15 设随机样本 X_1，…，X_n 取自 $\mathrm{n}(\theta,a\theta^2)$总体，其中 a 为已知常数，$\theta>0$.

(a) 证明参数空间不包含二维开集；

(b) 证明统计量 $T=(\overline{X},S^2)$是 θ 的充分统计量，但 $\mathrm{n}(\theta,a\theta^2)$分布类不是完全的.

6.19

6.19 随机变量 X 取值 0，1，2，且分布为下列之一：

	$P(X=0)$	$P(X=0)$	$P(X=0)$	
分布 1	p	$3p$	$1-4p$	$(0<p<\frac{1}{4})$
分布 2	p	p^2	$1-p-p^2$	$(0<p<\frac{1}{2})$

判断 X 的上述分布类是否是完全的.

6.21

6.21 设随机变量 X 的概率密度函数为：

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

(a) X 是否是完全充分统计量？

(b) $|X|$ 是否是完全充分统计量？

(c) $f(x|\theta)$是否属于指数族？