概率论

第二册 数理统计

第二分册

复旦大学 编

内 容 提 要

全书分为三册：概率论基础、数理统计、随机过程，材料比较丰富，并配备一定数量的习题。

数理统计分为两个分册出版，这是第二分册，介绍估计、假设检验的理论，以及序贯分析、非参数统计和多元分析的基本知识，可供综合大学和高等师范院校数学和其他有关专业作为试用教材。

全书由卞国瑞、吴立德、李贤平、汪嘉冈（依姓氏笔划为序）编写，本册使用说明详见第一册序言。

高等学校试用教材

概 率 论

第二册 数理统计

第二分册

复旦大学 编

*

高等教育出版社出版

新华书店北京发行所发行

江苏六合印刷厂印装

开本787×1092 1/32 印张8 2/16 字数195,000

1979年8月第1版 1986年3月第7次印刷

印数 118,401—127,950

书号13010·0381 定价1.15元

目　　录

附表

第五章 点 估 计

§1.3 中讨论了两种常用的参数估计方法，从那里已经看到，对于同一个未知参数可以构造出不同的估计量. 此时应选取哪一个为好呢？这就要求进一步研究点估计的性质，以帮助我们决定估计量的选取和提供求得优良估计的方法.

§1. 最小方差无偏估计

一、最小方差与无偏性

显然我们希望一个"好的"估计，在待估计的参数值的左右摆动且尽可能靠近. 但是估计量是子样的函数，它是随机变量，对不同的观察结果，所求得的参数估计值是不同的. 因而我们确定一个估计量的好坏不能根据某一次试验的结果来衡量，而应从总体而言，也就是说，在多次观察中我们希望好的估计都集中在真参数值的周围波动. 若 T_1, T_2 是未知参数 θ 的两个估计，当 θ 为真参数值时，它们的分布密度函数 $g_1(x)$ 和 $g_2(x)$ 如图 1 所示. 由图 1 直观可见 T_1 比 T_2 更集中在未知参数 θ 的周围，所以作为 θ 的估计，T_1 比 T_2 好. 用数学语言表达是对任一常数 $c>0$，有

$$P_\theta\{|T_1-\theta|<c\}$$
$$\geqslant P_\theta\{|T_2-\theta|<c\}$$

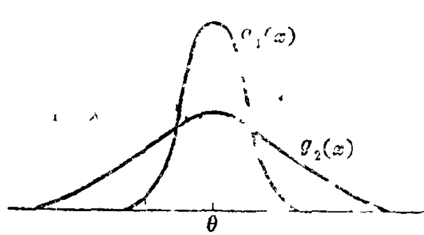

图 1

这里直接用抽样分布描写"集中"的概念是合理的，但要在此意义下寻找最佳估计很困难. 因此我们经常采用均方误差 $E_\theta(T-\theta)^2$ 作为衡量"集中"程度的尺度. 这就导致了最小方差无

偏估计问题.

定义 1 设母体$\mathbb{X}$具有分布族$\{F(x;\ \theta),\ \theta\in\Theta\}$，设$X=(X_1,\ \cdots,\ X_n)$是从该母体中抽取的一个子样. $T(X)$是θ的一个估计. 如果它满足下面的关系式

$$E_\theta(T)=\theta,\ \text{对一切}\ \theta\in\Theta \tag{1}$$

则称$T(X)$是参数θ的一个**无偏估计**.

如果θ是向量参数，定义1也同样适用，但要求(1)式中的T也是向量，且其维数与θ的维数相同.

如果有一列θ的估计$T_n\triangleq T_n(X_1,\ \cdots,\ X_n)$满足下面的关系式

$$\lim_{n\to\infty}E_\theta(T_n)=\theta,\ \text{对一切}\ \theta\in\Theta \tag{2}$$

则称T_n是θ的**渐近无偏估计**.

对θ的任一函数$g(\theta)$，如果对$g(\theta)$存在有无偏估计，也就是说，存在估计量T使得下式成立，

$$E_\theta(T)=g(\theta),\ \text{对一切}\ \theta\in\Theta$$

则称$g(\theta)$为**可估计函数**.

一个估计如果不是无偏的就称这估计是有偏的，且称函数

$$b(\theta,\ T)=E_\theta(T)-\theta \tag{3}$$

为估计T的偏.

［例 1］ 设$X_1,\ X_2,\ \cdots,\ X_n$是从具有有限数学期望的母体中抽取的一个子样，因为由定理1.2.1知道

$$E\overline{X}=\mu$$

所以子样均值$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$是母体均值$EX=\mu$的一个无偏估计. 但$\overline{X}^2$不是$\mu^2$的无偏估计，因为$E\overline{X}^2=\frac{1}{n}\sigma^2+\mu^2$. $\overline{X}^2$作为μ^2的估计，它的偏是$b(\mu^2,\ \overline{X}^2)=\frac{1}{n}\sigma^2$.

如果母体的方差 σ^2 存在，则由定理1.2.2知道

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{n}{n-1}S_n^2 \tag{4}$$

是方差 σ^2 的无偏估计．

但是，S_n^* 一般不是 σ 的无偏估计．如设 $X_1, \cdots, X_n$ 是从母体 $N(\mu, \sigma^2)$ 中抽取的一个子样．由定理2 4.1知道 $(n-1)S_n^{*2}/\sigma^2=nS_n^2/\sigma^2$ 服从 $\chi^2(n-1)$ 分布，所以

$$E_\sigma(\sqrt{n-1}\,S_n^*/\sigma)=\int_0^\infty\frac{\sqrt{x}}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}e^{-\frac{x}{2}}x^{\frac{n-1}{2}-1}dx$$

$$=\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)=\frac{\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$E_\sigma(S_n^*)=\sigma\left[\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]\neq\sigma$$

S_n^* 作为 σ 的估计，它的偏是

$$b(\sigma, S_n^*)=\sigma\left[\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}-1\right]$$

由此可见，如果 T 是参数 θ 的无偏估计，除了 f 是线性函数以外，并不能推出 T 的函数 $f(T)$ 也是参数 $f(\theta)$ 的无偏估计．

有时候，无偏估计可以不存在(见习题1, 2)；有时候，一个无偏估计可以有明显弊病(见习题2)；而有时候，对同一个参数，可以有很多无偏估计．如

[例2]　设母体 $\mathbb{X}$ 具有普阿松分布 $P(\lambda)$，$X_1, X_2, \cdots, X_n$ 是一子样．由于 $E\mathbb{X}=D(\mathbb{X})=\lambda$，所以 $\overline{X}$ 和 S_n^{*2} 都是 λ 的无偏估计．因此，对任 $0\leqslant\alpha\leqslant1$，$\alpha\overline{X}+(1-\alpha)S_n^{*2}$ 也是 λ 的无偏估计．

由此可见，仅有无偏性要求是不够的. 事实上它仅反映估计是在真参数 θ 的周围波动，而没有反映出"集中"的程度. 为此提出.

定义 2 设 $\theta_0\in\Theta$, 记

$$U(\theta_0)\triangleq\{T: E_\theta(T)=\theta,\ 对一切\ \theta\in\Theta;\ E_{\theta_0}(T^2)<\infty\} \tag{5}$$

如果对 $T_0\in U(\theta_0)$ 成立下式:

$$D_{\theta_0}(T_0)\leqslant D_{\theta_0}(T),\ 对一切\ T\in U(\theta_0) \tag{6}$$

则称 T_0 是在 θ_0 的**局部最小方差无偏估计.**

定义 3 记

$$U\triangleq\{T: E_\theta T=\theta,\ E_\theta T^2<\infty,\quad 对一切\ \theta\in\Theta\ 成立\} \tag{7}$$

即 U 是带有有限方差的 θ 的全体无偏估计所组成的集合. 如果对 $T_0\in U$ 成立下式,

$$D_\theta(T_0)\leqslant D_\theta(T),\ 对一切\ \theta\in\Theta\ 和\ T\in U \tag{8}$$

则称 T_0 是 θ 的**一致最小方差无偏估计.**

下面的定理给出了一个估计是一致最小方差无偏估计的充分必要条件.

定理 1 设由(7)式所定义的 U 是非空集合，记 U_0 是 0 的具有有限方差的无偏估计的全体所组成的集合，即

$$U_0\triangleq\{\nu: E_\theta\nu=0,\ E_\theta\nu^2<\infty,\ 对一切\ \theta\in\Theta\} \tag{9}$$

则 $T_0\in U$ 是 θ 的一个一致最小方差无偏估计的充要条件是

$$E_\theta(\nu T_0)=0.\ 对一切\ \theta\in\Theta\ 和\ \nu\in U_0 \tag{10}$$

[证明] 必要性:

用反证法，设 $T_0\in U$ 是 θ 的一致最小方差无偏估计，而条件(10)不满足. 则存在 $\nu_0\in U_0$ 和 $\theta_0\in\Theta$, 使得

$$E_{\theta_0}(T_0\nu_0)\neq 0$$

因为 $E_\theta\nu_0=0$, 所以对一切 λ, $T_\lambda\triangleq T_0-\lambda\nu_0\in U_0$, 计算

$$E_{\theta_0}T_\lambda^2=E_{\theta_0}(T_0^2)+\lambda^2E_{\theta_0}\nu_0^2-2\lambda E_{\theta_0}(T_0\nu_0)$$

由于 $$E_{\theta_0}(T_0\nu_0)\neq 0$$

一定能找到 $\lambda_0\left(例\ \lambda_0=\dfrac{E_{\theta_0}(T_0\nu_0)}{E_{\theta_0}(\nu_0^2)}\right)$, 使得

$$E_{\theta_0}(T_{\lambda_0}^2)<E_{\theta_0}(T_0^2) \tag{11}$$

所以 $$D_{\theta_0}(T_{\lambda_0})<D_{\theta_0}(T_0)$$

这与 T_0 是一致最小方差无偏估计的约定相矛盾，因此条件(10)是必要的.

充分性:

设对某个 $T_0\in U$, (10) 式成立. 则对任一 $T\in U$, 显然 $T-T_0\in U_0$, 所以

$$E_\theta\{T_0(T-T_0)\}=0, \quad 对一切\ \theta\in\Theta$$

由许瓦兹不等式得

$$E_\theta(T_0^2)=E_\theta(T_0T)\leqslant[E_\theta(T_0^2)]^{1/2}[E_\theta(T^2)]^{1/2}, \quad 对一切\ \theta\in\Theta$$

所以

$$E_\theta(T_0^2)\leqslant E_\theta(T^2), \quad 对一切\ \theta\in\Theta \tag{12}$$

又因为对一切 $\theta\in\Theta$, $E_\theta T=E_\theta T_0=\theta$, 故由(12)式得到

$$D_\theta(T_0)\leqslant D_\theta(T), \quad 对一切\ \theta\in\Theta$$

由于 T 是 U 中任一估计. 所以 T_0 是 θ 的一致最小方差无偏估计.

定理 2 设 U 是由(7)式定义的 θ 的非空无偏估计类, 则对 θ 最多存在一个一致最小方差无偏估计.

[证明] 如果 $T, T_0\in U$ 是 θ 的两个一致最小方差无偏估计, 则

$$\begin{aligned} E_\theta T_0=E_\theta T=\theta, \\ D_\theta(T_0)=D_\theta(T), \end{aligned} \quad 对一切\ \theta\in\Theta \tag{13}$$

因此

$$E_\theta(T-T_0)=0, \quad 对一切\ \theta\in\Theta \tag{14}$$

即 $$T-T_0\in U_0$$

所以

$$E_\theta\{T_0(T-T_0)\}=0, \quad 对一切\ \theta\in\Theta$$
$$E_\theta\{T(T-T_0)\}=0, \quad 对一切\ \theta\in\Theta \tag{15}$$

由(15)式得到

$$E_\theta(T-T_0)^2=E_\theta\{T(T-T_0)\}-E_\theta\{T_0(T-T_0)\}$$
$$=0, \quad 对一切\ \theta\in\Theta$$

由上式立即得到

$$P_\theta(T=T_0)=1, \quad 对一切\ \theta\in\Theta$$

这表明 θ 的一致最小方差无偏估计在概率 1 相等的意义下是唯一的.

系 1 设 T_1 和 T_2 分别是 $g_1(\theta)$ 和 $g_2(\theta)$ 的一致最小方差无偏估计，则 aT_1+bT_2 是 $ag_1(\theta)+bg_2(\theta)$ 的一致最小方差无偏估计. 其中 a、b 是固定常数.

它是定理 1 的直接推论，证明留作练习.

[例 3] 设母体 X 服从正态分布 $N(\mu, \sigma^2)$，其中 μ 和 σ^2 是未知参数. 设 $X=(X_1, \cdots, X_n)$ 是一简单子样，欲求 μ 和 σ^2 的一致最小方差无偏估计.

运用定理 1. 设 $\nu\in U_0$，则有

$$\int\cdots\int\nu\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}dx\equiv0 \tag{16}$$

上式关于 μ 微分，得

$$\int\cdots\int\nu\left(\sum x_i\right)\exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right\}dx\equiv0$$

由定理 1 得到，$\sum\limits_{i=1}^n X_i$ 是它的数学期望 $E\left(\sum\limits_{i=1}^n X_i\right)=n\mu$ 的一致最小方差无偏估计.

(16)式关于 μ 微分两次得

$$\int\cdots\int\nu\left(\sum x_i\right)^2\exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right\}dx=0$$

(16)式关于σ^2微分得

$$\int\cdots\int \nu \sum(x_i-\mu)^2 \exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right\}dx=0$$

由此可得

$$\int\cdots\int \nu \sum(x_i-\bar{x})^2 \exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right\}dx=0$$

所以由定理 1, $\sum(X_i-\bar{X})^2$ 是它的数学期望 $(n-1)\sigma^2$ 的一致最小方差无偏估计. 再运用系 1 得到 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是 μ 的一致最小方差无偏估计; $S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是 σ^2 的一致最小方差无偏估计.

如果估计 T 是子样的线性函数, 即 T 可以表示为 $T=\sum_{i=1}^{n}a_iX_i$, 其中 $a_1, \cdots, a_n$ 是固定常数, 则称 T 为**线性估计**. 类似地可以定义, 如果 T 是线性估计, 且满足无偏性条件(1)式, 则 T 称为**线性无偏估计**; 如果 U_L 表示 θ 的具有有限方差的线性无偏估计的全体所组成的集合, 而对 $T_0\in U_L$, 有

$$D_\theta(T_0)\leqslant D_\theta(T), \quad 对一切\ \theta\in\Theta\ 和\ T\in U_L \tag{17}$$

则称 T 为 θ 的**一致最小方差线性无偏估计**.

若用 U_L 代替 U, U_{L_0} 代替 U_0, 对于一致最小方差线性无偏估计, 类似的定理 1、定理 2、系 1 都成立. 在例 3 中, $\bar{X}$ 也是 μ 的一致最小方差线性无偏估计.

二、有效估计

上面讨论了一致最小方差无偏估计, 一个自然的问题是: 无偏估计的方差是否可以任意小? 如果不可以任意小那么它的下限是什么? 这个下限能否达到, 为此引入下面的定理.

定理 3 (罗(C. R. Rao)-克拉美(H. Cramer)不等式)

设 Θ 是实数轴上的一个开区间, $\{f(x;\theta), \theta\in\Theta\}$ 是母体 X 的

分布密度族；$X\triangleq(X_1, \cdots, X_n)$是从该母体中抽取的一个简单随机子样；$T(X)$ 是 $g(\theta)$ 的一个无偏估计．且满足条件：

(1) 集合 $S_\theta\triangleq\{x: f(x;\theta)\neq 0\}$ 与 θ 无关；

(2) $g'(\theta)$ 和 $\dfrac{\partial f(x;\theta)}{\partial\theta}$ 存在，且

$$\frac{\partial}{\partial\theta}\int f(x,\theta)dx=\int\frac{\partial f(x;\theta)}{\partial\theta}dx,\quad 对一切\ \theta\in\Theta$$

$$\frac{\partial}{\partial\theta}\int_{R_n}T(x)L(x;\theta)dx=\int_{R_n}T(x)\frac{\partial L(x;\theta)}{\partial\theta}dx,\quad 对一切\ \theta\in\Theta$$

其中 $L(x;\theta)=\prod\limits_{i=1}^{n}f(x_i;\theta)$ 是子样 $X=(X_1, \cdots, X_n)$ 的联合分布密度；

(3) $I(\theta)\triangleq E_\theta\left(\dfrac{\partial\log f(X;\theta)}{\partial\theta}\right)^2>0$，则

$$[g'(\theta)]^2\leqslant D_\theta(T)\cdot nI(\theta),\quad 对一切\ \theta\in\Theta \tag{18}$$

如果上式中等号成立，则存在 $c(\theta)\neq 0$ 使得下式成立，

$$\frac{\partial\log L_\theta(X)}{\partial\theta}=c(\theta)\cdot[T(X)-g(\theta)],\ a.\ s. \tag{19}$$

如果 $g'(\theta_0)\neq 0$，则有

$$D_{\theta_0}(T)\geqslant\frac{[g'(\theta_0)]^2}{nI(\theta_0)} \tag{20}$$

特别当 $g(\theta)=\theta$ 时，上式可简化为

$$D_\theta(T)\geqslant\frac{1}{nI(\theta)} \tag{21}$$

［证明］ 若 $I(\theta)=+\infty$ 或 $D_\theta(T)=+\infty$，则 (18) 式显然成立，故可假设 $I(\theta)<\infty$ 和 $D_\theta(T)<\infty$．

由条件(1)和(2)得

$$\int_{-\infty}^{\infty}\frac{\partial\log f(x;\theta)}{\partial\theta}f(x;\theta)dx=\int_{-\infty}^{\infty}\frac{\partial f(x;\theta)}{\partial\theta}dx$$

$$=\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f(x;\theta)dx=0$$

所以 $$\int_{R_n}\frac{\partial\log L(x;\theta)}{\partial\theta}L(x;\theta)dx=0,\quad 对一切\ \theta\in\Theta$$

$$\int_{R_n}g(\theta)\frac{\partial\log L(x,\theta)}{\partial\theta}L(x;\theta)dx=0$$

又 $$g'(\theta)=\frac{\partial}{\partial\theta}\int_{R_n}T(x)L(x;\theta)dx=\int_{R_n}T(x)\frac{\partial L(x;\theta)}{\partial\theta}dx$$
$$=\int_{R_n}T(x)\frac{\partial\log L(x;\theta)}{\partial\theta}L(x;\theta)dx$$
$$=\int_{R_n}[T(x)-g(\theta)]\frac{\partial\log L(x;\theta)}{\partial\theta}L(x;\theta)dx$$
$$对一切\ \theta\in\Theta$$

由许瓦兹不等式得

$$[g'(\theta)]^2=\left\{\int_{R_n}[T(x)-g(\theta)]\frac{\partial\log L(x;\theta)}{\partial\theta}L(x;\theta)dx\right\}^2$$
$$\leqslant\int_{R_n}[T(x)-g(\theta)]^2L(x;\theta)dx$$
$$\times\int_{R_n}\left[\frac{\partial\log L(x;\theta)}{\partial\theta}\right]^2L(x;\theta)dx$$
$$=D_\theta(T)\cdot nI(\theta),\quad 对一切\ \theta\in\Theta$$

如果等号成立，则存在 $c(\theta)\neq0$ 使得

$$\frac{\partial\log L(X;\theta)}{\partial\theta}=c(\theta)[T(X)-g(\theta)],\ a.s.$$

最后 (20), (21) 两式是显然的，定理证毕.

对于离散型随机变量，只要将定理 3 条件中的密度函数用概率分布代替，积分号用求和号代替，结论完全成立. 留给读者作练习.

对于多个参数的情形，有完全类似的结果.

定理 4 设母体 X 具有分布密度族 $\{f(x;\theta),\ \theta\in\Theta\}$，$\theta\triangleq(\theta_1,\cdots,\theta_k)^\tau$. $X=(X_1,\cdots,X_n)$ 是一简单随机子样，$T(X)\triangleq(T_1(X),\cdots,T_r(X))^\tau$ 是 $g(\theta)\triangleq(g_1(\theta),\cdots,g_r(\theta))^\tau$ 的一个无

偏估计，即

$$E_\theta T(X)=g(\theta),\ 对一切\ \theta\in\Theta$$

且满足条件：

(1) $\frac{\partial}{\partial\theta_i}\int L(x;\theta)dx=\int\frac{\partial}{\partial\theta_i}L(x;\theta)dx$，对一切 $\theta\in\Theta$ 和 $i=1,\cdots,k$

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j}\int L(x;\ \theta)dx=\int\frac{\partial^2}{\partial\theta_i\partial\theta_j}L(x;\ \theta)dx$$

对一切 $\theta\in\Theta$ 和 $i,\ j=1,\ \cdots,\ k$

其中
$$L(x;\ \theta)\triangleq\prod_{i=1}^n f(x_i;\ \theta)$$

(2) $\frac{\partial g_i}{\partial\theta_j}$ 对 $i=1,\ \cdots,\ r;\ j=1,\ \cdots,\ k$ 存在，且

$$\frac{\partial g_i}{\partial\theta_j}=\frac{\partial}{\partial\theta_j}\int T_iL(x;\ \theta)dx=\int T_i\frac{\partial L(x;\ \theta)}{\partial\theta_j}dx$$

对一切 $\theta\in\Theta,\ i=1,\ \cdots,\ r,\ j=1,\ \cdots,\ k$

记
$$I(\theta)\triangleq-E_\theta\begin{pmatrix}\frac{\partial^2\log f}{\partial^2\theta_1} & \frac{\partial^2\log f}{\partial\theta_1\partial\theta_2}\cdots & \frac{\partial^2\log f}{\partial\theta_1\partial\theta_k}\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{\partial^2\log f}{\partial\theta_k\partial\theta_1} & \frac{\partial^2\log f}{\partial\theta_k\partial\theta_2}\cdots & \frac{\partial^2\log f}{\partial^2\theta_k}\end{pmatrix}_{k\times k}$$

$$\Sigma_\theta(T)\triangleq D_\theta(T)\triangleq(\mathrm{cov}(T_i,\ T_j))_{r\times r}$$

$$\Delta\triangleq\begin{pmatrix}\frac{\partial g_1}{\partial\theta_1}\cdots & \frac{\partial g_1}{\partial\theta_k}\\ \cdots\cdots & \cdots\cdots\\ \frac{\partial g_r}{\partial\theta_1}\cdots & \frac{\partial g_r}{\partial\theta_k}\end{pmatrix}_{r\times k}$$

则有

$$\Sigma_\theta(T)\geqslant\frac{1}{n}\Delta I^{-1}(\theta)\Delta^\tau,\ 对一切\ \theta\in\Theta \tag{22}$$

即 $\Sigma_\theta(T)-\frac{1}{n}\Delta I^{-1}(\theta)\Delta^\tau$ 是非负定矩阵①。

① 定理 4 的证明及有关罗-克拉美不等式的深入讨论可参阅 C. R. Rao, Linear Statistical Inference and Its Applications, 第二版，第五章。

[例 4] 设母体 X 具有贝努里分布 $b(1, p)$，$p\in\Theta=(0, 1)$，$(X_1, \cdots, X_n)$ 是一子样，试求 p 的无偏估计的方差下界。

容易验证定理 3 的条件满足，且

$$I(p)=\sum_{x=0}^{1}\left(\frac{\partial \log f_p(x)}{\partial p}\right)^2 f_p(x)=\frac{1}{p(1-p)}$$

$$f_p(x)=\begin{cases} p, & \text{当 } x=1 \\ 1-p, & \text{当 } x=0 \end{cases}$$

所以方差下限是

$$\frac{1}{nI(p)}=\frac{p(1-p)}{n}$$

大家知道 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{\nu}{n}$（$\nu$ 表示"1"发生的频数）是 p 的无偏估计，而 $D_p(\overline{X})=\frac{1}{n}p(1-p)$ 达到罗-克拉美不等式的下界，所以 $\frac{\nu}{n}$ 是 p 的一致最小方差无偏估计.

[例 5] 设 $X_1, \cdots, X_n$ 是取自正态母体 $\{N(\mu, \sigma^2), -\infty<\mu<\infty, \sigma^2>0\}$ 的一个子样，考虑 μ 和 σ^2 的无偏估计的方差下限。

容易计算

$$I(\mu)=\int_{-\infty}^{\infty}\left(\frac{x-\mu}{\sigma^2}\right)^2\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}dx=\frac{1}{\sigma^2}$$

因此 μ 的无偏估计的方差下限是 σ^2/n，而

$$D(\overline{X})=\frac{1}{n}\sigma^2$$

这表示子样均值 $\overline{X}$ 达到了方差下限，所以它是 μ 的一致最小方差无偏估计.

同样地

$$I(\sigma^2)=\int_{-\infty}^{\infty}\left[\frac{(x-\mu)^2}{2\sigma^4}-\frac{1}{2\sigma^2}\right]^2\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx$$
$$=\frac{1}{2\sigma^4}$$

因此 σ^2 的无偏估计的方差下限是 $\frac{2}{n}\sigma^4$.

由例 1 知 S_n^{*2} 是 σ^2 的无偏估计，用定理 1.2.1 得

$$D(S_n^{*2})=\frac{2}{n-1}\sigma^4$$

因此
$$D(S_n^{*2})=\frac{2}{n-1}\sigma^4>\frac{2}{n}\sigma^4=\frac{1}{nI(\sigma^2)}$$

例 3 曾指出，S_n^{*2} 是 σ^2 的一致最小方差无偏估计，这说明罗-克拉美不等式中的方差下限不一定可达到.

下面我们引进有效性的概念.

定义 4 如果 T 是参数 θ 的一个无偏估计，它的方差达到 (21) 式给出的下界，那么称 T 是**有效估计**.

对于 θ 的任一无偏估计 T，记

$$e_\theta(T)\triangleq\frac{1}{nI(\theta)}\Big/D_\theta(T) \tag{23}$$

我们称 $e_\theta(T)$ 为无偏估计 T 的**有效率**. 显然，对于满足定理 3 条件的无偏估计 T，其有效率满足 $0\leqslant e_\theta(T)\leqslant 1$. 如果无偏估计 T 是有效估计，则它的有效率 $e_\theta(T)=1$.

粗略地说，对于一个无偏估计，估计的方差愈小，这个估计取到接近于它的期望(即 θ)的值就愈频繁. 因而有效率愈大，方差愈小，"未知参数的估计值在它真值附近"的概率愈大.

由例 4 看到，对于贝努里概型，频率是其对应的概率的有效估计. 由例 5 可见，对于正态 $N(\mu, \sigma^2)$ 母体，子样均值 $\overline{X}$ 是母体均数 μ 的有效估计，修正后的子样方差 S_n^{*2} 是 σ^2 的一致最小方差无偏估计，但不是 σ^2 的有效估计，对于 S_n^{*2}，它的有效率是

$$e_{\sigma^2}(S_n^{*2})=\frac{2\sigma^4}{n}\Big/\frac{2\sigma^4}{n-1}=\frac{n-1}{n}<1$$

显然当 $n\to\infty$ 时有 $e_{\sigma^2}(S_n^{*2})\to 1$，即有效率渐近于 1，以后称具有这种性质的估计为渐近有效估计.

定义 5 设 $T_n(X_1, \cdots, X_n)$ 是 θ 的一列无偏估计，若

$$\lim_{n\to\infty}\frac{\dfrac{1}{nI(\theta)}}{D_\theta(T_n)}=\lim_{n\to\infty}e(T_n)=e_0 \tag{24}$$

则称 e_0 是 T_n 的渐近有效率，而当 $e_0=1$ 时，我们称 T_n 是 θ 的渐近有效估计。

§2. 充分性和完备性

充分性和完备性是数理统计中两个很重要的基本概念，它也有助于我们寻找一致最小方差无偏估计。

一、充分统计

[例 1] 设母体 X 服从正态分布 $N(\mu, 1)$，要求估计均值 μ。在例 2.5 中已经指出子样均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是 μ 的一个有效估计。大家知道，一个容量为 n 的子样 $(X_1, X_2, \cdots, X_n)$ 包含了 n 个值。而现在估计 μ 时，对这 n 个资料进行了“压缩”和“精简”，仅用一个单值函数 $\overline{X}$。这里自然地会提出一个问题，这种压缩是否合理？也就是说，$\overline{X}$ 是否已经包含了子样关于 μ 的全部信息？用子样的个别值或其它形式能否更多地知道 μ 呢？

下面就回答这个问题。我们知道 $(X_1, \cdots, X_n)$ 是 n 维空间中的一个点，而当已知 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=y$ 时，子样 $(X_1, \cdots, X_n)$ 是 $n-1$ 维空间中的点，因而上面提出的问题就归结为假若已经知道 $\overline{X}=y$ 时，进一步知道子样 $(X_1, \cdots, X_n)$ 在这 $n-1$ 维空间中的位置，关于 μ 能否获得更多的信息？为说明简便，对 $n=2$ 时进行考虑。此时子样空间 (X_1, X_2) 可以用平面来表示，$\overline{X}=\frac{1}{2}(X_1+X_2)=y$ 表示该平面中的一条直线（见图 1）。在这平面上，(X_1, X_2) 服从于二元正态分布，它的分布密度函数是