

概率论与数理统计 (Fall 2024)

习题课讲义

2024 年 12 月 4 日

聂嵘琢, 祝国隽

B01GB006Y-02/Note12

1. 令总体 X 的概率函数为

$$P(X = -1) = \frac{1-\theta}{2}, P(X = 0) = \frac{1}{2}, P(X = 1) = \frac{\theta}{2}, 0 < \theta < 1.$$

记 X_1, X_2, \dots, X_n iid $\sim X$.

(1) 求 θ 的无偏估计;

(2) 求 θ 的无偏估计的方差的 C-R 不等式下界.

Ans: (1) 注意到 $E(X) = -\frac{1-\theta}{2} + \frac{\theta}{2} = \theta - \frac{1}{2}$, 则可构造

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^n X_i + \frac{1}{2}$$

且

$$E(\hat{\theta}_1) = \frac{1}{n} \sum_{i=1}^n E(X_i) + \frac{1}{2} = \theta$$

很多同学这里用最大似然估计, 也可以, 但相对来说复杂很多

(2) 下面先求 θ 的信息量.

$$\begin{aligned} \ln f(x, \theta) &= \frac{1}{2} (x^2 - x) \ln \left(\frac{1-\theta}{2} \right) - (1-x^2) \ln 2 + \frac{1}{2} (x^2 + x) \ln \left(\frac{\theta}{2} \right) \\ \frac{\partial \ln f(x, \theta)}{\partial \theta} &= -\frac{1}{2} (x^2 - x) \frac{1}{1-\theta} + \frac{1}{2} (x^2 + x) \frac{1}{\theta} \\ I(\theta) &= E \left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 = \left(-\frac{1}{1-\theta} \right)^2 \frac{1-\theta}{2} + 0 \times \frac{1}{2} + \left(\frac{1}{\theta} \right)^2 \frac{\theta}{2} \\ &= \frac{1}{2} \left(\frac{1}{1-\theta} + \frac{1}{\theta} \right)^2 = \frac{1}{2\theta(1-\theta)} \end{aligned}$$

$$\theta \text{ 的无偏估计的方差的 C-R 下界为 } \frac{1}{nI(\theta)} = \frac{2\theta(1-\theta)}{n}.$$

2. 令总体 X 的密度函数为

$$f(x, \theta) = \begin{cases} \theta x^{\theta-1}, & 0 < x < 1, \\ 0, & \text{其他} \end{cases}$$

其中 $\theta > 0, g(\theta) = 1/\theta$. 记 $X_1, X_2, \dots, X_n \text{ iid } \sim X$.

(1) 求 $g(\theta)$ 的矩估计 $\hat{g}_1(\theta)$ 和最大似然估计 $\hat{g}_2(\theta)$;

(2) $\hat{g}_2(\theta)$ 是否为无偏估计?

(3) 判断 $\hat{g}_2(\theta)$ 的方差是否达到 C-R 不等式下界.

Ans: (1). 由于 $E(X_1) = \int_0^1 x\theta x^{\theta-1}dx = \frac{\theta}{\theta+1}x^{\theta+1}\Big|_0^1 = \frac{\theta}{\theta+1} = \frac{1}{1+g(\theta)}$, 令 $\bar{X} = \sum_{i=1}^n X_i$ 为样本均值, 令

$$\frac{1}{1+g(\theta)} = \bar{X},$$

则 $\hat{g}_1(\theta) = \frac{1}{\bar{X}} - 1$ 为 $g(\theta)$ 的矩估计.

考虑到, 似然函数为 $L(\theta) = \prod_{i=1}^n (\theta X_i^{\theta-1}) = \theta^n \prod_{i=1}^n X_i^{\theta-1}$, 对数似然函数为

$$\ln L(\theta) = n \ln \theta + (\theta - 1) \sum_{i=1}^n \ln X_i = -n \ln g(\theta) + \left(\frac{1}{g(\theta)} - 1 \right) \sum_{i=1}^n \ln X_i,$$

关于 $g(\theta)$ 求导并令其为 0, 得

$$\frac{\partial \ln L(\theta)}{\partial g(\theta)} = -\frac{n}{g(\theta)} - \frac{1}{g^2(\theta)} \sum_{i=1}^n \ln X_i = 0$$

解得 $\hat{g}_2(\theta) = -\frac{1}{n} \sum_{i=1}^n \ln X_i$, 是 $g(\theta)$ 的极大似然估计.

(2). 令 $Y = -\ln X$, 则

$$P(Y < y) = P(-\ln X < y) = P(X > e^{-y}) = \int_{e^{-y}}^1 \theta x^{\theta-1} dx = 1 - e^{-\theta y},$$

从而 Y 服从指数分布 $\text{Exp}(\theta)$, 也即 $Y \sim \text{Exp}(\theta) = \text{Gamma}(1, \theta)$, 根据 Gamma 分布的可加性, 有 $n\hat{g}_2(\theta) \sim \text{Gamma}(n, \theta)$, 从而

$$E(n\hat{g}_2(\theta)) = \frac{n}{\theta}, E(\hat{g}_2(\theta)) = \frac{1}{\theta} = g(\theta),$$

则 $\hat{g}_2(\theta)$ 是无偏的.

(3). 下面求 θ 的 Fisher 信息量, 注意到 $\ln f(x; \theta) = \ln \theta + (\theta - 1) \ln X$, $\frac{\partial \ln f(x; \theta)}{\partial \theta} = \frac{1}{\theta} + \ln X$, $\frac{\partial^2 \ln f(x; \theta)}{\partial \theta^2} = -\frac{1}{\theta^2}$, 于是

$$I(\theta) = -E\left(\frac{\partial^2 \ln f(x; \theta)}{\partial \theta^2}\right) = \frac{1}{\theta^2}.$$

而 $g'(\theta) = -\frac{1}{\theta^2}$, 于是 $g(\theta)$ 的任一无偏估计的方差的 C-R 下界为

$$\frac{[g'(\theta)]^2}{nI(\theta)} = \frac{1}{n\theta^2}$$

针对 $g(\theta)$ 的极大似然估计 $\hat{g}_2(\theta)$, 由于 $\hat{g}_2(\theta) \sim \text{Gamma}(n, n\theta)$, 那么

$$\text{Var}(\hat{g}_2(\theta)) = \frac{n}{(n\theta)^2} = \frac{1}{n\theta^2}$$

由于其达到了 C-R 下界, 所以 $\hat{g}_2(\theta)$ 是 $g(\theta)$ 的 UMVUE.

3. 令总体 $X \sim B(1, p)$, $0 < p < 1$. 记 X_1, X_2, \dots, X_n iid $\sim X$, $n \geq 2$. 求:

(1) 充分完备统计量;

(2) $p(1-p)$ 的 UMVUE.

Ans: (1). $(X_1, X_2, \dots, X_n)^\top$ 的联合密度函数为

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} \\ &= p^{\sum_{i=1}^n x_i} (1-p)^{\sum_{i=1}^n (1-x_i)} \\ &= \exp \left\{ \sum_{i=1}^n x_i \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1-p) \right\}. \end{aligned}$$

则 $T = \sum_{i=1}^n x_i$ 是充分统计量, 同时它又是完备统计量. 从而 $T = \sum_{i=1}^n x_i$ 是充分完备统计量.

(2). 令 $\hat{g} = \hat{g}(X_1, X_2, \dots, X_n) = \frac{1}{2} I_{\{X_1+X_2=1\}}, E\hat{g} = p(1-p),$

求 $E(\hat{g} | \sum_{i=1}^n X_i = t) = \frac{n\bar{X}(1-\bar{X})}{n-1}.$

$$\begin{aligned} E(\hat{g} | T = t) &= \frac{1}{2} P\left(X_1 + X_2 = 1 \mid \sum_{i=1}^n X_i = t\right) \\ &= \frac{1}{2} \frac{P(X_1 + X_2 = 1, \sum_{i=1}^n X_i = t-1)}{\sum_{i=1}^n X_i = t} \\ &= \frac{1}{2} \frac{P(X_1 + X_2 = 1) P(\sum_{i=1}^n X_i = t-1)}{\sum_{i=1}^n X_i = t} \\ &= \frac{t(n-t)}{n(n-1)}. \end{aligned}$$

4. 令总体 X 的分布函数

$$F(x, \theta) = \begin{cases} 1 - e^{-x^2/\theta}, & x \geq 0, \\ 0, & \text{其他}, \end{cases}$$

其中 $\theta > 0$. 记 $X_1, X_2, \dots, X_n \text{ iid } \sim X$.

- (1) 求 X 的均值和方差;
- (2) 求 θ 的矩估计和最大似然估计;
- (3) 最大似然估计是 θ 的相合估计吗?

Ans: (1) 密度函数为

$$f(x, \theta) = \begin{cases} \frac{2x}{\theta} e^{-\frac{x^2}{\theta}}, & x > 0 \\ 0, & \text{其他} \end{cases}.$$

$$\begin{aligned} E(X_i) &= \int_0^\infty x \frac{2x}{\theta} e^{-\frac{x^2}{\theta}} dx = \int_0^\infty x d\left(-e^{-\frac{x^2}{\theta}}\right) = -xe^{-\frac{x^2}{\theta}} \Big|_0^\infty + \int_0^\infty e^{-\frac{x^2}{\theta}} dx \\ &= \sqrt{\theta\pi} \int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{\theta/2}} e^{-\frac{x^2}{2 \times \theta/2}} dx = \frac{\sqrt{\theta\pi}}{2}. \\ E(X_i^2) &= \int_0^\infty x^2 \frac{2x}{\theta} e^{-\frac{x^2}{\theta}} dx = \int_0^\infty x^2 d\left(-e^{-\frac{x^2}{\theta}}\right) = \int_0^\infty 2xe^{-\frac{x^2}{\theta}} dx \\ &= \theta \int_0^\infty d\left(-e^{-\frac{x^2}{\theta}}\right) = \theta. \end{aligned}$$

则 $\text{Var}(X_i) = EX_i^2 - (EX_i)^2 = \theta - \frac{\pi}{4}\theta$, 其中 $\theta \geq \frac{\pi}{4}$.

(2) 记 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 为样本均值. 令 $\bar{x} = \frac{\sqrt{\theta\pi}}{2}$, 则 $\hat{\theta}_1 = \frac{4\bar{x}^2}{\pi}$ 为 θ 的矩估计. 似然函数为

$$L(\theta) = \prod_{i=1}^n \frac{2x_i}{\theta} e^{-\frac{x_i^2}{\theta}} I_{\{x_i > 0\}} = \left[\prod_{i=1}^n (2x_i) \right] \frac{1}{\theta^n} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i^2}.$$

对数似然为

$$l(\theta) = \sum_{i=1}^n \ln(2x_i) - n \ln \theta - \frac{1}{\theta} \sum_{i=1}^n x_i^2.$$

似然方程为

$$\frac{\partial l(\theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i^2}{\theta^2} = 0.$$

解得 $\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^n x_i^2$ 为 θ 的极大似然估计.

(3) 由于 $E(\hat{\theta}_2) = E(X_1^2) = \theta$, $E(X_1^4) = \int_0^\infty x^4 \frac{2x}{\theta} e^{-\frac{x^2}{\theta}} dx = 2\theta^2$. $\text{Var}(\hat{\theta}_2) = \frac{1}{n} \text{Var}(X_1^2) = \frac{1}{n} [EX_1^4 - \theta^2] = \frac{\theta^2}{n} \rightarrow 0 (n \rightarrow \infty)$. 从而极大似然估计 $\hat{\theta}_2$ 是 θ 的相合估计.

5. 记 X_1, X_2, \dots, X_n iid $\sim U(0, \theta)$, $\theta > 0$, 次序统计量的最大值为 $X_{(n)}$. 求证: $X_{(n)}$ 是 θ 的相合估计.

Ans: 由于 $X_{(n)}$ 的密度函数为 $f(u) = n \left(\frac{u}{\theta}\right)^{n-1} \frac{1}{\theta}, 0 < u < \theta$. 且

$$E(X_{(n)}) = \int_0^\theta u f(u) du = \int_0^\theta \frac{n}{\theta^n} u^n du = \frac{n}{n+1} \theta \rightarrow \theta (n \rightarrow \infty)$$

$$E(X_{(n)}^2) = \int_0^\theta u^2 f(u) du = \int_0^\theta \frac{n}{\theta^n} u^{n+1} du = \frac{n}{n+2} \theta^2.$$

则

$$\begin{aligned} \text{Var}(X_{(n)}) &= E(X_{(n)}^2) - (E(X_{(n)}))^2 = \theta^2 \left[\frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2 \right] \\ &= \frac{n}{(n+1)^2(n+2)} \theta^2 \rightarrow 0 (n \rightarrow \infty) \end{aligned}$$

由定理 6.7. 有 $X_{(n)}$ 是 θ 的相合估计.

补充材料参考张伟平老师 (USTC) 的讲义, 关于 UMVUE, 主要是两种套路.

Lec6: 点估计(三)

张伟平

2011 年 3 月 24 日

1 一致最小方差无偏估计

一、引言及定义

设有一参数分布族 $\mathcal{F} = \{F_\theta, \theta \in \Theta\}$, 其中 Θ 为参数空间. 设 $g(\theta)$ 是定义在 Θ 上的函数, $\mathbf{X} = (X_1, \dots, X_n)$ 为自总体 F_θ 中抽取的简单样本, $\hat{g}(\mathbf{X}) = \hat{g}(X_1, \dots, X_n)$ 为 $g(\theta)$ 的一个估计量, 如何评价 $\hat{g}(\mathbf{X})$ 的优劣? 一般用 $\hat{g}(\mathbf{X}) - g(\theta)$ 作为其偏差, 为消除 $\hat{g}(\mathbf{X}) - g(\theta)$ 取值出现 “+,-” 可能抵消的影响, 一般用 $(\hat{g}(\mathbf{X}) - g(\theta))^2$ 来代替. 由于这个量是随机的, 将其平均, 即计算其均值, 以得到一个整体性的指标 $E_\theta(\hat{g}(\mathbf{X}) - g(\theta))^2$, 这就是估计量 $\hat{g}(\mathbf{X})$ 的均方误差.

定义 1. 设 $\hat{g}(\mathbf{X})$ 为 $g(\theta)$ 的估计, 则称 $E_\theta(\hat{g}(\mathbf{X}) - g(\theta))^2$ 为 $\hat{g}(\mathbf{X})$ 的均方误差 (Mean Square Error, 简记 MSE)

设 $\hat{g}_1(\mathbf{X})$ 和 $\hat{g}_2(\mathbf{X})$ 为 $g(\theta)$ 的两个不同的估计, 若

$$E_\theta(\hat{g}_1(\mathbf{X}) - g(\theta))^2 \leq E_\theta(\hat{g}_2(\mathbf{X}) - g(\theta))^2, \quad \text{对一切 } \theta \in \Theta,$$

则称在 MSE 准则下 $\hat{g}_1(\mathbf{X})$ 优于 $\hat{g}_2(\mathbf{X})$.

若存在 $\hat{g}^*(\mathbf{X})$, 使得对 $g(\theta)$ 的任一估计量 $\hat{g}(\mathbf{X})$, 都有

$$E_\theta(\hat{g}^*(\mathbf{X}) - g(\theta))^2 \leq E_\theta(\hat{g}(\mathbf{X}) - g(\theta))^2, \quad \text{对一切 } \theta \in \Theta,$$

则称 $\hat{g}^*(\mathbf{X})$ 为 $g(\theta)$ 的一致最小均方误差估计.

可惜的是, 一致最小均方误差估计常不存在. 解决这个问题的办法之一, 是把最优性准则放宽一些, 使适合这种最优性准则的估计一般能存在. 从直观上想, 在一个大的估计量的类中找一致最优的估计不存在, 把估计量的类缩小, 就有可能存在一致最优的估计量. 因此我们把估计类缩小为无偏估计类来考虑. 在无偏估计类中, 估计量的均方误差就变为其方差. 即当 $\hat{g}(\mathbf{X})$ 为 $g(\theta)$ 的无偏估计时, $MSE(\hat{g}(\mathbf{X})) = D_\theta(\hat{g}(\mathbf{X}))$, 此处 $D_\theta(\hat{g}(\mathbf{X}))$ 表示 $\hat{g}(\mathbf{X})$ 的方差.

存在这样的情形, 对参数 $g(\theta)$ 它的无偏估计不存在. 请看下例:

例 1. 设样本 $X \sim$ 二项分布 $b(n, p)$, n 已知而 p 未知. 令 $g(p) = 1/p$, 则参数 $g(p)$ 的无偏估计不存在.

证 采用反证法: 若不然, $g(p)$ 有无偏估计 $\hat{g}(X)$. 由于 X 只取 $0, 1, \dots, n$ 这些值, 令 $\hat{g}(X)$ 的取值用 $\hat{g}(i) = a_i$ 表示, $i = 0, 1, \dots, n$. 由 $\hat{g}(X)$ 的无偏性, 应有

$$E_p(\hat{g}(X)) = \sum_{i=0}^n a_i \binom{n}{i} p^i (1-p)^{n-i} = 1/p, \quad 0 < p < 1.$$

于是有

$$\sum_{i=0}^n a_i \binom{n}{i} p^{i+1} (1-p)^{n-i} - 1 = 0, \quad 0 < p < 1.$$

但上式左端是 p 的 $n+1$ 次多项式, 它最多在 $(0, 1)$ 区间有 $n+1$ 个实根, 可无偏性要求对 $(0, 1)$ 中的任一实数 p 上式都成立. 这个矛盾说明 $g(p) = 1/p$ 无的偏估计不存在.

今后我们把不存在无偏估计的参数除外. 参数的无偏估计若存在, 则此参数为可估参数; 若参数函数的无偏估计存在, 则称此函数为可估函数 (Estimable function). 因此可估函数的无偏估计类是非空的.

假如可估函数的无偏估计类中的无偏估计不止一个, 怎样比较它们的优劣? 故引入下列的定义.

定义 2. 设 $\mathcal{F} = \{F_\theta, \theta \in \Theta\}$ 是一个参数分布族, 其中 Θ 为参数空间, $g(\theta)$ 为定义在 Θ 上的可估函数. 设 $\hat{g}^*(\mathbf{X}) = \hat{g}^*(X_1, \dots, X_n)$ 为 $g(\theta)$ 的一个无偏估计, 若对 $g(\theta)$ 的任一无偏估计 $\hat{g}(\mathbf{X}) = \hat{g}(X_1, \dots, X_n)$, 都有

$$D_\theta(\hat{g}^*(\mathbf{X})) \leq D_\theta(\hat{g}(\mathbf{X})), \quad \text{对一切 } \theta \in \Theta,$$

则称 $\hat{g}^*(\mathbf{X})$ 是 $g(\theta)$ 的一致最小方差无偏估计 (Uniformly Minimum Variance Unbiased Estimation, 简记为 UMVUE).

对给定参数分布族, 如何寻找可估参数的 UMVUE 呢? 本节以下将介绍两种方法: 零无偏估计法和充分完全统计量法, 下一节的 Cramer-Rao 不等式法也是寻找 UMVUE 的一种方法.

在前面我们曾介绍过 Rao-Blackwell 定理, 这一定理提供了一个改进无偏估计的方法, 它在本节以下寻找 UMVUE 中, 起到简化问题的作用. 重新表述此定理如下

定理 1 (Rao-Blackwell). 设 $T = T(\mathbf{X})$ 是一个充分统计量, 而 $\hat{g}(\mathbf{X})$ 是 $g(\theta)$ 的一个无偏估计, 则

$$h(T) = E(\hat{g}(\mathbf{X})|T)$$

是 $g(\theta)$ 的无偏估计, 并且

$$D_\theta(h(T)) \leq D_\theta(\hat{g}(\mathbf{X})), \quad \text{一切 } \theta \in \Theta, \quad (1.1)$$

其中等号当且仅当 $P_\theta(\hat{g}(\mathbf{X}) = h(T)) = 1$, 即 $\hat{g}(\mathbf{X}) = h(T)$, a.e. P_θ 成立.

这个引理提供了一个改进无偏估计的方法, 即一个无偏估计 $\hat{g}(\mathbf{X})$ 对充分统计量 $T(\mathbf{X})$ 的条件期望 $E\{\hat{g}(\mathbf{X})|T\}$ 将能导出一个新的无偏估计, 且它的方差不会超过原估计量 $\hat{g}(\mathbf{X})$ 的方差. 若原估计 $\hat{g}(\mathbf{X})$ 不是 $T(\mathbf{X})$ 的函数, 则新的无偏估计 $E(\hat{g}(\mathbf{X})|T)$ 一定比原估计 $\hat{g}(\mathbf{X})$ 具有更小的方差. 这个定理还表明一致最小方差无偏估计一定是充分统计量的函数, 否则可以通过充分统计量构造出一个具有更小方差的无偏估计来.

例2. 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是从两点分布族 $\{b(1, p) : 0 < p < 1\}$ 中抽取的简单样本. 显然, X_1 是 p 的一个无偏估计, $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 是 p 的充分统计量, 试利用 $T = T(\mathbf{X})$ 构造一个具有比 X_1 方差更小的无偏估计.

解 由引理3.4.1可知, 容易构造 p 的一个无偏估计如下:

$$h(T) = E(X_1|T = t) = 1 \cdot P(X_1 = 1|T = t) + 0 \cdot P(X_1 = 0|T = t)$$

$$\begin{aligned}
&= \frac{P(X_1 = 1, T = t)}{P(T = t)} = \frac{P(X_1 = 1, X_2 + \cdots + X_n = t - 1)}{P(T = t)} \\
&= \frac{p \cdot \binom{n-1}{t-1} p^{t-1} (1-p)^{n-t}}{\binom{n}{t} p^t (1-p)^{n-t}} = \frac{t}{n} = \bar{x}.
\end{aligned}$$

显然样本均值 $h(T) = \bar{X}$ 的方差为 $p(1-p)/n$, 而 X_1 的方差为 $p(1-p)$, 当 $n \geq 2$ 时 \bar{X} 的方差更小.

二、零无偏估计法

本段介绍一个一般性的定理, 用以判断某一估计量是否为UMVUE.

定理 2. 设 $\hat{g}(\mathbf{X})$ 是 $g(\theta)$ 的一个无偏估计, $D_\theta(\hat{g}(\mathbf{X})) < \infty$, 对任何 $\theta \in \Theta$. 且对任何满足条件 “ $E_\theta l(\mathbf{X}) = 0$, 对一切 $\theta \in \Theta$ ” 的统计量 $l(\mathbf{X})$, 必有

$$Cov_\theta(\hat{g}(\mathbf{X}), l(\mathbf{X})) = E_\theta[\hat{g}(\mathbf{X}) \cdot l(\mathbf{X})] = 0, \quad \text{一切 } \theta \in \Theta, \quad (1.2)$$

则 $\hat{g}(\mathbf{X})$ 是 $g(\theta)$ 的UMVUE.

从形式上看, 条件 “ $E_\theta l(\mathbf{X}) = 0$, 对一切 $\theta \in \Theta$ ” 可解释为 “ $l(\mathbf{X})$ 是零的无偏估计”, 由此得到求UMVUE的方法之一的名称 “零无偏估计法”.

此定理还可进一步加强为: 设 $\hat{g}(\mathbf{X})$ 为 $g(\theta)$ 的一个无偏估计, $D_\theta(\hat{g}(\mathbf{X})) < \infty$, 一切 $\theta \in \Theta$, 则 $\hat{g}(\mathbf{X})$ 是 $g(\theta)$ 的UMVUE的充分必要条件是: 对任何满足条件 “ $E_\theta l(\mathbf{X}) = 0$, 对一切 $\theta \in \Theta$ ” 的统计量 $l(\mathbf{X})$, 必有(1.2)成立.

证 设 $\hat{g}_1(\mathbf{X})$ 为 $g(\theta)$ 的任一无偏估计. 记 $l(\mathbf{X}) = \hat{g}_1(\mathbf{X}) - \hat{g}(\mathbf{X})$ 为零的无偏估计, 由于(1.2)式成立, 因而

$$\begin{aligned}
D_\theta(\hat{g}_1(\mathbf{X})) &= D_\theta[\hat{g}(\mathbf{X}) + l(\mathbf{X})] \\
&= D_\theta(\hat{g}(\mathbf{X})) + D_\theta(l(\mathbf{X})) + 2Cov_\theta(\hat{g}(\mathbf{X}), l(\mathbf{X})) \\
&= D_\theta(\hat{g}(\mathbf{X})) + D_\theta(l(\mathbf{X})) \geq D_\theta(\hat{g}(\mathbf{X})).
\end{aligned}$$

另一方面, 如果 $\delta = \hat{g}(X)$ 为 $g(\theta)$ 的UMVUE, 则对任意的零无偏估计量 $l(x)$, 令

$$\delta' = \delta + \lambda l(X)$$

则 δ' 仍为 $g(\theta)$ 的无偏估计, 且

$$Var(\delta') = Var(\delta + \lambda l(X)) = \lambda^2 Var(l(X)) + 2\lambda Cov(\delta, l(X)) + Var(\delta) \geq Var(\delta).$$

对所有的 λ 成立. 等价于

$$\lambda^2 Var(l(X)) + 2\lambda Cov(\delta, l(X)) \geq 0, \quad \forall \lambda.$$

左边有两个根 $\lambda = 0$ 和 $\lambda = -Cov(\delta, l(X))/Var(l(X))$, 推出 $Cov(\delta, l(X)) = 0$.

这就证明了所要的结果.

从定理的内容看, 它是一个验证某个特定的估计量 $\hat{g}(\mathbf{X})$ 为UMVUE的方法. 至于这个特定的估计 $\hat{g}(\mathbf{X})$ 从何而来, 定理3.4.1不能提供任何帮助, 它不是UMVUE的构造性定理. $\hat{g}(\mathbf{X})$ 可以从直观的想法提出, 如由矩估计或极大似然估计等方法获得, 然后利用此定理验证它是否

为 $g(\theta)$ 的UMVUE. 条件(1.2)的验证也不容易, 因为零无偏估计很多. 下面的几个例子说明, 本章前面一些例子中提到的几个常用估计, 都可以用此法验证其为UMVUE.

例 求上例中 $g(p) = p$ 的UMVUE.

解 由上例已知 $T = \sum_{i=1}^n X_i$ 为充分统计量, $\hat{g}(\mathbf{X}) = T/n$, 故只要验证它满足定理的条件. 显然 $\hat{g}(\mathbf{X})$ 是 p 的无偏估计. 且 $D_\theta(\hat{g}(\mathbf{X})) = p(1-p)/n < \infty$, 对一切 $0 < p < 1$. 现设 $l = l(T)$ 为任一零无偏估计, 并记 $a_i = l(i)$, $i = 0, 1, 2, \dots, n$, 则因 $T \sim b(n, p)$, 故有

$$E_p l(T) = \sum_{i=0}^n a_i \binom{n}{i} p^i (1-p)^{n-i} = 0, \quad 0 < p < 1.$$

约去因子 $(1-p)^n$, 并记 $\varphi = p/(1-p)$ (φ 取值 $(0, \infty)$), 将上式改写为

$$\sum_{i=0}^n a_i \binom{n}{i} \varphi^i = 0, \quad \text{一切 } 0 < \varphi < \infty.$$

上式左边是 φ 的多项式, 要使其为0, 必有 $a_i \binom{n}{i} = 0$, 即 $a_i = 0$, $i = 1, 2, \dots, n$. 故 $l(T)$ 在其定义域中处处为0, 因而有 $l(T) \equiv 0$. 从而有

$$\text{Cov}_p(\hat{g}, l(T)) = E(\hat{g} \cdot l(T)) = 0.$$

即定理条件成立, 故 $\hat{g}(\mathbf{X}) = \bar{X} = T/n$ 为 p 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从指数分布 $EP(\lambda)$ 中抽取的简单样本, 求总体均值 $g(\lambda) = 1/\lambda$ 的UMVUE.

解 由于在指数分布族中 $T = \sum_{i=1}^n X_i$ 为 λ 的充分统计量, 则 $T \sim \Gamma(n, \lambda)$, 其密度函数为

$$\phi(t, \lambda) = \begin{cases} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} & \text{当 } t > 0 \\ 0 & \text{当 } t \leq 0, \end{cases}$$

其中 $\lambda > 0$. 取 $\hat{g}(\mathbf{X}) = T/n$, 显然 $E(\hat{g}(\mathbf{X})) = 1/\lambda$, 即 $\hat{g}(\mathbf{X})$ 为 $g(\lambda) = 1/\lambda$ 的无偏估计, 且 $D_\lambda(\hat{g}(\mathbf{X})) = 1/(n\lambda^2) < \infty$. 现设 $l = l(T)$ 为任一零无偏估计, 故有

$$El(T) = \int_0^\infty l(t) \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt = 0,$$

即 $\int_0^\infty l(t) t^{n-1} e^{-\lambda t} dt = 0$. 两边对 λ 求导得

$$\int_0^\infty l(t) t^n e^{-\lambda t} dt = 0,$$

此式等价于 $E_\lambda[T/n \cdot l(T)] = E_\lambda(\hat{g} \cdot l(T)) = \text{Cov}_\lambda(\hat{g} \cdot l) = 0$, 即定理条件成立. 因此 $\hat{g}(\mathbf{X}) = T/n$ 为 $g(\lambda) = 1/\lambda$ 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从均匀分布 $U(0, \theta)$ 中抽取的简单样本, 求 θ 的UMVUE.

解 由 $T = T(\mathbf{X}) = X_{(n)}$ 是参数 θ 的充分统计量, 又知 $\hat{g}(\mathbf{X}) = \frac{n+1}{n}T$ 是 θ 的无偏估计, 且 $D_\theta(\hat{g}(T)) = \frac{1}{n(n+2)}\theta^2 < \infty$. 现设 $l(T)$ 为任一零无偏估计, T 的密度函数如(??)所示, 因此有

$$E_\theta l(T) = \int_0^\theta l(t) \cdot \frac{nt^{n-1}}{\theta^n} dt = 0, \quad \text{一切 } \theta > 0,$$

于是有

$$\int_0^\theta l(t)t^{n-1}dt = 0, \quad \text{一切 } \theta > 0.$$

将上式两边对 θ 求导得 $l(\theta)\theta^{n-1} = 0$, 一切 $\theta > 0$. 故有 $l(\theta) \equiv 0$ 对一切 $\theta > 0$. 可见 $Cov(\hat{g}, l(T)) = E(\hat{g} \cdot l(T)) = 0$, 即定理条件成立. 因此 $\hat{g}(\mathbf{X}) = \frac{n+1}{n}X_{(n)}$ 为 $g(\theta) = \theta$ 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从正态分布 $N(a, \sigma^2)$ 中抽取的简单随机样本, 求 a 和 σ^2 的UMVUE.

解 由 $T = (T_1, T_2)$ 为 $\theta = (a, \sigma^2)$ 的充分统计量. 其中 $T_1 = \bar{X}$, $T_2 = \sum_{i=1}^n (X_i - \bar{X})^2$, 又 T_1 和 T_2 独立, 且 $T_1 \sim N(a, \sigma^2/n)$, $T_2/\sigma^2 \sim \chi_{n-1}^2$. 因此 (T_1, T_2) 的联合密度为

$$f(t_1, t_2) = \frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} e^{-\frac{n(t_1-a)^2}{2\sigma^2}} \cdot \left(2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right) \sigma^{n-1}\right)^{-1} t_2^{\frac{n-3}{2}} e^{-\frac{t_2}{2\sigma^2}},$$

当 $-\infty < t_1 < +\infty$, $t_2 > 0$; 其它处为 0. (1.3)

先考虑 a 的UMVUE, 令 $\hat{g}_1(\mathbf{X}) = T_1$, 显然 $\hat{g}_1(\mathbf{X})$ 为 $g_1(\theta) = a$ 的无偏估计, 且 $D_\theta(\hat{g}_1(\mathbf{X})) = \sigma^2/n < \infty$. 现设 $l(T) = l(T_1, T_2)$ 为任一零无偏估计, 则有

$$E_\theta(l(T)) = \int_0^\infty \int_{-\infty}^\infty l(t_1, t_2) f_\theta(t_1, t_2) dt_1 dt_2 = 0,$$

此处 $-\infty < a < \infty$, $\sigma > 0$. 将上式两边对 a 求导数, 得

$$\int_0^\infty \int_{-\infty}^\infty l(t_1, t_2) (t_1 - a) \cdot t_2^{\frac{n-3}{2}} \exp\left\{-\frac{1}{2\sigma^2} [n(t_1 - a)^2 + t_2]\right\} dt_1 dt_2 = 0,$$

按(1.3), 此即

$$Cov_\theta(\hat{g}_1, l(T)) = E_\theta(\hat{g}_1 \cdot l(T_1, T_2)) = 0, \quad -\infty < a < +\infty, \quad \sigma > 0,$$

故定理3.4.1的条件满足. 所以 $\hat{g}_1(\mathbf{X}) = T_1$ 为 $g_1(\theta) = a$ 的UMVUE.

同理可验证 $T_2 = S^2$ 为 $g_2(\theta) = \sigma^2$ 的UMVUE, 这一验证留为练习.

三、充分完全统计量法

下列定理给出的求UMVUE的方法, 即充分完全统计量法是由E.L. Lehmann和H. Scheffe给出的, 完全统计量的概念也是由他们在1950年提出的.

定理 (Lehmann-Scheffe定理) 设 $T(\mathbf{X})$ 为一个充分完全统计量, 若 $\hat{g}(T(\mathbf{X}))$ 为 $g(\theta)$ 的一个无偏估计, 则 $\hat{g}(T(\mathbf{X}))$ 是 $g(\theta)$ 的唯一的UMVUE (唯一性是在这样的意义下: 若 \hat{g} 和 \hat{g}_1 都是 $g(\theta)$ 的UMVUE, 则 $P_\theta(\hat{g} \neq \hat{g}_1) = 0$, 对一切 $\theta \in \Theta$).

证 先证唯一性. 设 $\hat{g}_1(T(\mathbf{X}))$ 为 $g(\theta)$ 的任一无偏估计, 令 $\delta(T(\mathbf{X})) = \hat{g}(T(\mathbf{X})) - \hat{g}_1(T(\mathbf{X}))$, 则 $E_\theta \delta(T(\mathbf{X})) = E_\theta \hat{g}(T(\mathbf{X})) - E_\theta \hat{g}_1(T(\mathbf{X})) = 0$, 对一切 $\theta \in \Theta$. 由 $T(\mathbf{X})$ 为完全统计量, 可知 $\delta(T(\mathbf{X})) = 0$, a.e. P_θ . 即 $\hat{g}(T(\mathbf{X})) = \hat{g}_1(T(\mathbf{X}))$, a.e. P_θ , 故唯一性成立.

设 $\varphi(\mathbf{X})$ 为 $g(\theta)$ 的任一无偏估计. 令 $h(T(\mathbf{X})) = E(\varphi(\mathbf{X})|T)$, 由 $T(\mathbf{X})$ 为充分统计量, 故知 $h(T(\mathbf{X}))$ 与 θ 无关, 因此是统计量, 可知

$$E_\theta(h(T(\mathbf{X}))) = g(\theta), \quad \text{一切 } \theta \in \Theta,$$

$$D_\theta(h(T(\mathbf{X}))) \leq D_\theta(\varphi(\mathbf{X})), \quad \text{一切 } \theta \in \Theta.$$

由唯一性得 $\hat{g}(T(\mathbf{X})) = h(T(\mathbf{X}))$, $a.e. P_\theta$ 故有

$$D_\theta(\hat{g}(T(\mathbf{X}))) \leq D_\theta(\varphi(\mathbf{X})), \quad \text{一切 } \theta \in \Theta,$$

所以 $\hat{g}(T(\mathbf{X}))$ 为 $g(\theta)$ 的UMVUE, 且唯一.

推论 设样本 $\mathbf{X} = (X_1, \dots, X_n)$ 的分布为指数族

$$f(\mathbf{x}, \theta) = C(\theta) \exp \left\{ \sum_{j=1}^k \theta_j T_j(\mathbf{x}) \right\} h(\mathbf{x}), \quad \theta = (\theta_1, \dots, \theta_k) \in \Theta.$$

令 $T(\mathbf{X}) = (T_1(\mathbf{X}), \dots, T_k(\mathbf{X}))$, 若自然参数空间 Θ 作为 R_k 的子集有内点, 且 $h(T(\mathbf{X}))$ 为 $g(\theta)$ 的无偏估计, 则 $h(T(\mathbf{X}))$ 为 $g(\theta)$ 的唯一的UMVUE.

证 在推论的条件下, 由指数族的性质可知 $T(\mathbf{X})$ 为充分完全统计量. 故由 *Lehmann-Scheffé* 定理 (简记为L-S定理), 得知 $h(T(\mathbf{X}))$ 为 $g(\theta)$ 的唯一的UMVUE.

例 证明两点分布的 p 的无偏估计 $\hat{g}(\mathbf{X}) = T/n = \bar{X}$ 为 p 的UMVUE.

证 由因子分解定理可知 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 为两点分布 $b(1, p)$ 中参数 p 的充分统计量, 由例2.8.1知 $T(\mathbf{X})$ 也是完全统计量, 故 $\hat{g}(\mathbf{X}) = T/n = \bar{X}$ 是充分完全统计量 $T(\mathbf{X})$ 的函数, 且 $E_p \hat{g}(\mathbf{X}) = p$, 对 $0 < p < 1$. 因此由L-S定理可知 $\hat{g}(\mathbf{X})$ 为 p 的唯一的UMVUE.

例 在上例中, 已知 $T = \sum_{i=1}^n X_i$ 服从二项分布 $b(n, p)$, 且 $T(\mathbf{X})$ 为充分完全统计量, 求 $g(p) = p(1-p)$ 的UMVUE.

解 设 $\delta(T)$ 为 $g(p) = p(1-p)$ 的一个无偏估计, 要导出 $\delta(T)$ 的表达式. 按无偏估计的定义及 $T \sim b(n, p)$, 可得

$$\sum_{t=0}^n \binom{n}{t} \delta(t) p^t (1-p)^{n-t} = p(1-p), \quad \text{一切 } 0 < p < 1.$$

令 $\rho = p/(1-p)$, 故有 $p = \rho/(1+\rho)$, $1-p = 1/(1+\rho)$, 将它们代入上式得

$$\sum_{t=0}^n \binom{n}{t} \delta(t) \rho^t = \rho(1+\rho)^{n-2}, \quad 0 < \rho < \infty.$$

将 $\rho(1+\rho)^{n-2}$ 展开得 $\rho(1+\rho) = \sum_{l=0}^{n-2} \binom{n-2}{l} \rho^{l+1} = \sum_{t=1}^{n-1} \binom{n-2}{t-1} \rho^t$, 将其代入上式右边得

$$\sum_{t=0}^n \binom{n}{t} \delta(t) \rho^t = \sum_{t=1}^{n-1} \binom{n-2}{t-1} \rho^t, \quad 0 < \rho < \infty.$$

上式两边为 ρ 的多项式, 比较其系数得

$$\begin{aligned} \delta(t) &= 0, \quad \text{当 } t = 0, n; \\ \delta(t) &= \binom{n-2}{t-1} / \binom{n}{t} = \frac{t(n-t)}{n(n-1)}, \quad \text{当 } t = 1, 2, \dots, n-1. \end{aligned}$$

综合上述两式得

$$\delta(T) = \frac{T(n-T)}{n(n-1)}, \quad t = 0, 1, \dots, n$$

为 $g(p) = p(1-p)$ 的无偏估计, 它又是充分完全统计量 $T = \sum_{i=1}^n X_i$ 的函数, 由L-S定理可知 $\delta(T)$ 为 $g(p)$ 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从Poisson分布 $P(\lambda)$ 中抽取的简单随机样本, 求 (1) $g_1(\lambda) = \lambda$; (2) $g_2(\lambda) = \lambda^r$, $r > 0$ 为自然数; (3) $g_3(\lambda) = P_\lambda(X_1 = x)$ 的UMVUE.

解 由§2.7和§2.8可知 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 为关于Poisson分布的充分完全统计量.

(1) 令 $\hat{g}_1(T) = T(\mathbf{X})/n$, $E(\hat{g}_1(T)) = E(\bar{X}) = \lambda$, 故 $\hat{g}_1(T)$ 是分完全统计量 T 的函数, 且是 λ 的无偏估计, 故由L-S定理可知 $\hat{g}_1(T)$ 是 λ 的UMVUE.

(2) 由于 $T(\mathbf{X}) = \sum_{i=1}^n X_i \sim P(n\lambda)$, 令 $\delta(T)$ 为 $g_2(\lambda) = \lambda^r$ 的无偏估计, 故有 $E_\lambda \delta(T) = g_2(\lambda)$, 即

$$\sum_{t=0}^{\infty} \delta(t) \frac{e^{-n\lambda} (n\lambda)^t}{t!} = \lambda^r.$$

此式等价于

$$\sum_{t=0}^{\infty} \delta(t) \frac{n^t \lambda^t}{t!} = \lambda^r e^{n\lambda}.$$

将上式右边作展开得

$$\lambda^r e^{n\lambda} = \sum_{l=0}^{\infty} \frac{n^l \lambda^{l+r}}{l!} = \sum_{t=r}^{\infty} \frac{n^{t-r} \lambda^t}{(t-r)!}.$$

将其代入上式右边得

$$\sum_{t=0}^{\infty} \delta(t) \frac{n^t \lambda^t}{t!} = \sum_{t=r}^{\infty} \frac{n^{t-r} \lambda^t}{(t-r)!}.$$

上述等式两边是 λ 的幂级数, 比较其系数得

$$\begin{aligned} \delta(t) &= 0, \quad \text{当 } t = 0, 1, \dots, r-1, \\ \delta(t) &= \frac{t! n^{t-r}}{(t-r)! n^t} = \frac{t(t-1) \cdots (t-r+1)}{n^r}, \quad \text{当 } t = r, r+1, \dots \end{aligned}$$

综合上述两式得

$$\delta(T) = \frac{T(T-1) \cdots (T-r+1)}{n^r}, \quad T = 0, 1, 2, \dots$$

为 $g_2(\lambda) = \lambda^r$ 的无偏估计, $\delta(T)$ 是充分完全统计量 T 的函数, 故由L-S定理可知 $\delta(T)$ 为 $g_2(\lambda)$ 的UMVUE.

(3) 由 $P_\lambda(X_1 = x) = e^{-\lambda} \lambda^x / x!$, 可见它是参数 λ 的函数, 故可用 $g_3(\lambda)$ 表示. 令 $\varphi(X_1) = I_{[X_1=x]}$, 则 $E_\lambda[\varphi(X_1)] = P_\lambda(X_1 = x)$. 因此 $\varphi(X_1)$ 为 $g_3(\lambda)$ 的无偏估计, 注意到 $T = \sum_{i=1}^n X_i \sim P(n\lambda)$ 和 $\sum_{i=2}^n X_i \sim P((n-1)\lambda)$, 故有

$$\begin{aligned} \delta_1(T) &= \delta_1(T(\mathbf{X})) = E(\varphi(X_1) | T = t) = \frac{P_\lambda(X_1 = x, T = t)}{P_\lambda(T = t)} \\ &= \frac{P_\lambda(X_1 = x) P_\lambda(X_2 + \cdots + X_n = t - x)}{P_\lambda(X_1 + \cdots + X_n = t)} = \frac{(n-1)^{t-x} t!}{n^t (t-x)! x!} \\ &= \binom{t}{x} \frac{(n-1)^{t-x}}{n^t} = \binom{t}{x} \left(\frac{1}{n}\right)^x \left(1 - \frac{1}{n}\right)^{t-x}, \quad t \geq x \end{aligned}$$

$\delta_1(T)$ 为 $g_3(\lambda)$ 的无偏估计, 它又是充分完全统计量 $T(\mathbf{X})$ 的函数, 所以

$$\delta_1(T(\mathbf{X})) = \binom{T}{x} \left(\frac{1}{n}\right)^x \left(1 - \frac{1}{n}\right)^{T-x}$$

为 $g_3(\lambda)$ 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从指数分布 $EP(\lambda)$ 中抽取的简单随机样本, 求 $g(\lambda) = \lambda$ 的UMVUE.

解 由 $T(\mathbf{X}) = \sum_{i=1}^n X_i$ 的充分完全统计量, 以及 $T(\mathbf{X}) \sim \Gamma(n, \lambda)$, 即参数为 n 和 λ 的Gamma分布, 故有

$$E\left(\frac{1}{T}\right) = \int_0^\infty \frac{1}{t} \cdot \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt = \frac{\lambda}{n-1}.$$

因此 $\hat{g}(T(\mathbf{X})) = (n-1)/T(\mathbf{X})$ 为 λ 的无偏估计, 由L-S定理可知它是 λ 的UMVUE.

例 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为从正态分布 $N(a, \sigma^2)$ 中抽取的简单随机样本, 记 $\theta = (a, \sigma^2)$.
(1)求 a 和 σ^2 的UMVUE, (2)求 $g(\theta) = \sigma^r$ 的UMVUE.

解 由 $T(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}))$, 其中 $T_1(\mathbf{X}) = \bar{X}$, $T_2(\mathbf{X}) = \sum_{i=1}^n (X_i - \bar{X})^2$, 为充分完全统计量.

(1)由于 $\hat{g}_1(\mathbf{X}) = \bar{X} = T_1$ 和 $\hat{g}_2(\mathbf{X}) = T_2/(n-1)$ 分别为 a 和 σ^2 的无偏估计, 它们又是充分完全统计量, 故由L-S定理可知它们分别是 a 和 σ^2 的UMVUE.

(2)由于 $T_2/\sigma^2 \sim \chi_{n-1}^2$, 故 σ^r 的无偏估计与 T_2 的幂函数有关. 先计算下式:

$$\begin{aligned} E\left(\frac{T_2}{\sigma^2}\right)^{r/2} &= \frac{1}{\sigma^r} E\left(T_2^{r/2}\right) = \int_0^\infty t^{r/2} \cdot \frac{1}{2^{n-1} \Gamma\left(\frac{n-1}{2}\right)} t^{\frac{n-1}{2}-1} e^{-\frac{t}{2}} dt \\ &= \frac{2^{r/2} \Gamma\left(\frac{n+r-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \triangleq \frac{1}{K_{n-1, r}}. \end{aligned}$$

由上式可知

$$E(K_{n-1, r} \cdot T_2^{r/2}) = \sigma^r.$$

因此估计量

$$\hat{g}_3(T(\mathbf{X})) = K_{n-1, r} T_2^{r/2} = \frac{\Gamma\left(\frac{n-1}{2}\right)}{2^{r/2} \Gamma\left(\frac{n+r-1}{2}\right)} T^{r/2}$$

是 σ^r 的无偏估计, 又是充分完全统计量 $T = (T_1, T_2)$ 的函数, 故由L-S定理可知它是 $g(\theta) = \sigma^r$ 的UMVUE.

例 用Lehmann-Scheffe定理再考虑均匀分布 $U(0, \theta)$, $\theta > 0$, 中的参数 θ 的UMVUE.

解 由 $T(\mathbf{X}) = \max(X_1, \dots, X_n) = X_{(n)}$ 为充分完全统计量, 和 $\hat{g}(T(\mathbf{X})) = \frac{n+1}{n} T(\mathbf{X})$ 为 θ 的无偏估计, 故由L-S定理立得 $\hat{g}(T(\mathbf{X}))$ 为 θ 的UMVUE. 可见此处证明要简单的多.

例 考虑均匀分布 $U(0, \theta)$, $\theta > 1$, 中的参数 θ 的UMVUE.

解 由因子分解定理知道 $T(\mathbf{X}) = \max(X_1, \dots, X_n) = X_{(n)}$ 为充分统计量, 但它并不为完全统计量. 实际上对任意满足 $El(T) = 0$, 即

$$\int_0^1 l(t) t^{n-1} dt + \int_1^\theta l(t) t^{n-1} dt = 0 \quad (*)$$

因此可取

$$l_0(t) = \begin{cases} (n+1)t - n, & 0 < t < 1 \\ 0, & 1 < t < \theta \end{cases}$$

因而存在非恒为零的函数 $l(T) = l_0(T)$ 满足 $El(T) = 0$.所以 T 不是完全统计量, 从而不能利用Lehmann-Scheffe定理验证基于 T 的无偏估计为UMVUE, 但是我们可以使用零无偏方法. 注意到欲使 $Eg(T)l(T) = 0$, 即

$$\int_0^1 g(t)l(t)t^{n-1}dt + \int_1^\theta g(t)l(t)t^{n-1}dt = 0$$

结合(*)式从而可取

$$g(t) = \begin{cases} c, & 0 < t < 1 \\ bt, & 1 < t < \theta \end{cases}$$

由无偏性有

$$Eg(T) = \int_0^1 cf_T(t)dt + \int_1^\theta btf_T(t)dt = \frac{c}{\theta^n} + \frac{b}{\theta^n} \frac{n}{n+1}(\theta^{n+1} - 1) = \theta$$

从而可取

$$c = 1, \quad b = \frac{n+1}{n}$$

因此

$$g(T) = \begin{cases} 1, & 0 < T < 1 \\ \frac{n+1}{n}T, & 1 < T < \theta \end{cases}$$

为参数 θ 的UMVUE.

注 可以证明统计量

$$\tilde{T} = \begin{cases} 1, & 0 < T < 1 \\ T, & 1 < T < \theta \end{cases}$$

为完全统计量.

2 Cramer-Rao不等式

一、引言

Cramer-Rao不等式(简称C-R不等式)是判别一个无偏估计量是否为UMVUE的方法之一. 这一方法的思想如下: 设 \mathcal{U}_g 是 $g(\theta)$ 的一切无偏估计构成的类. \mathcal{U}_g 中的估计量的方差有一个下界, 这个下界称为C-R下界. 因此, 如果 $g(\theta)$ 的一个无偏估计 \hat{g} 的方差达到这个下界, 则 \hat{g} 就是 $g(\theta)$ 的一个UMVUE, 当然样本分布族和 \hat{g} 要满足一定的正则条件. 这个不等式是由C.R. Rao和H. Cramer在1945和1946年分别证明的. 以后一些统计学者将条件作了一些改进和精确化, 但结果的基本形式并无重大变化.

这一方法的缺陷是: 由于C-R不等式确定的下界常比真下界为小. 在一些场合, 虽然 $g(\theta)$ 的UMVUE \hat{g} 存在, 但其方差大于C-R下界. 在这一情况下, 用C-R不等式就无法判定 $g(\theta)$ 的UMVUE存在. 因此这一方法的适用范围不广. C-R不等式除了用于判别 $g(\theta)$ 的UMVUE之外, 它在数理统计理论上还有其它的用处, 如估计的效率和有效估计的概念以及Fisher信息量都与之有关.

C-R不等式成立需要样本分布族满足一些正则条件, 适合这些条件的分布族称为C-R正则分布族, 下面给出其定义.