Chpt.7 Statistical Inference: Parameter Estimation

第七章 参数估计

习题：矩估计

设总体X的概率密度为

$$f(x;\theta)=\begin{cases} e^{-(x-\theta)}, & x \geqslant \theta \\ 0, & x<\theta \end{cases}$$

$(X_1,\cdots,X_n)$ 是取自总体X的样本，则参数θ的矩估计量为 ______

解：

$$EX=\int_0^{+\infty} x e^{-(x-\theta)} dx = e^{\theta}\cdot\int_{\theta}^{+\infty} x e^{-x} dx$$

$$= e^{\theta}\cdot\left(-x e^{-x}-e^{-x}\right)\Big|_{\theta}^{+\infty}=\theta+1$$

则有$\theta = \mathrm{EX}-1$，故$\hat{\theta} = \overline{\mathrm{X}}-1=\frac{1}{\mathrm{n}}\sum_{i=1}^{n} X_i-1$

习题：最大似然估计

设射手的命中率为p，在向同一目标的80次射击中，命中75次，则p的最大似然估计值为 ________

记

$$X = \begin{cases} 1, & 射击命中目标 \\ 0, & 射击没有命中 \end{cases}$$

则 X 服从两点分布，

X	1	0
P	p	1-p

既有 $P(X=x) = p^x(1-p)^{1-x}, x = 0,1$

似然函数为

$$L(p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

$$Ln\ L = \left(\sum_{i=1}^{n} x_i\right)\ln p + (n - \sum_{i=1}^{n} x_i)\ln(1-p)$$

$$\frac{dLn\ L}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$$

解得$\hat{p}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$，故 p 的最大似然估计值为

$$\frac{75}{80} = \frac{15}{16}$$

习题：矩估计和最大似然估计的对比

设总体X的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中$0<\theta<\frac{1}{2}$是未知参数，利用总体X的如下样本值3,1,3,0,3,1,2,3
求θ的矩估计值和最大似然估计值。

分析 参数的最大似然估计值，就是对给定的观测值$(x_1,\cdots,x_n)$，选取$\hat{\theta}$，使得似然函数L(θ)达到最大，本题中总体X是离散型随机变量，根据其分布确定似然函数L(θ)是求解的关键

习题：矩估计和最大似然估计的对比

解 (1)EX=0× θ^2+1×2θ(1-θ)+2× θ^2+3× $(1-2\theta)$ = 3-4θ

$$\theta = \frac{1}{4}(3\text{-EX}),\quad \hat{\theta}_{ME}=\frac{1}{4}(3-\overline{X})$$

故θ的矩阵估计值为

$$\hat{\theta}_{ME}=\frac{1}{4}(3-\bar{x})=\frac{1}{4}(3-2)=\frac{1}{4}$$

(2)对给定的样本值(3,1,3,0,3,1,2,3)，似然函数为

$$L(\theta)=\prod_{i=1}^{8}P\{X_i=x_i\}=P\{X=0\}\cdot\{P(X=1)\}^2\cdot\ P\{X=2\}\cdot\{P(X=3)\}^4$$

$$=\theta^2\cdot[2\theta(1-\theta)]^2\cdot\theta^2\cdot(1-2\theta)^4=4\theta^6(1-\theta)^2(1-2\theta)^4$$

$$\mathrm{Ln}\,L=\ln 4+6\ln\theta+2\ln(1-\theta)+4\ln(1-2\theta)$$

$$\frac{d\ln L}{d\theta}=\frac{6}{\theta}-\frac{2}{1-\theta}-\frac{8}{1-2\theta}=0$$

即有$12\theta^2-14\theta+3=0$

得$\hat{\theta}_{MLE}=\frac{7-\sqrt{13}}{12}$ ($\hat{\theta}=\frac{7+\sqrt{13}}{12}$ 舍去)

回46 回53

例题8：关于矩估计和最大似然估计的对比

设总体X~U[0,θ]，其中$\theta > 0$为未知参数，(X_1，…，X_n)是取自总体X的样本，求θ的矩估计量与最大似然估计量.

例题8：关于矩估计和最大似然估计的对比

设总体X~U[0,θ]，其中$\theta > 0$为未知参数，(X_1，…，X_n)是取自总体X的样本，求θ的矩估计量与最大似然估计量.

解 (1)EX=$\frac{\theta}{2}$，θ=2EX，$\hat{\theta}_{ME}=2\overline{X}$；

（2）X~$f(x;\theta)=\begin{cases} \frac{1}{\theta} & 0 \le x \le \theta \\ 0 & 其他 \end{cases}$

似然函数为

$$L(\theta) = \frac{1}{\theta^n}, 0 \leqslant x_i \leqslant \theta$$

$$\mathrm{Ln}\, L = -n \ln \theta$$

似然方程为

$$\frac{\mathrm{dLn}\, L}{\mathrm{d}\theta} = -\frac{n}{\theta} = 0$$

例题8：关于矩估计和最大似然估计的对比

虽然似然方程无解（由于参数θ>0）,注意到$\frac{dLnL}{d\theta}$<0，故$Ln(L)$关于θ单调递减，欲使$Ln(L)$达到最大，θ应最小，但是$0 \le x_i \le \theta$，故θ的最小值为$\max\limits_{1\le i\le n} x_i$，从而$\theta$的最大似然估计为$\hat{\theta}_{\mathrm{M}L\mathrm{E}}=\max\limits_{1\le i\le n} X_i$

注 本题目说明，当似然函数单调时，不能通过似然函数方程求解最大似然估计（似然方程无解）。此时应根据最大似然估计的概念，直接将似然函数的最大值点作为未知参数的最大似然估计。

[3] 衡量估计值优劣的标准

同一参数可能存在不同的估计量和估计值

比如：方差 σ^2 的估计量有

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

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

不同的估计量，哪个好，哪个差？

这是估计量的评选问题，需要建立一些标准

[3] 衡量估计值优劣的标准

1.无偏性

$X_1, X_2, \cdots, X_n$ 是总体X的样本

$\theta \in \Theta$ 是包含在总体X中的未知参数

若估计量 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 的数学期望 $E(\hat{\theta})$ 存在且等于未知参数 θ，即

$$E(\hat{\theta}) = \theta$$

则称 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 为 θ 的无偏估计量.

此时，用 $\hat{\theta}(x_1, \cdots, x_n)$ 代替 θ 不含系统误差.

估计量的无偏性是说对于某些样本值，由这一估计量得到的估计值相对于真值来说偏大，有些则偏小. 反复将这一估计量使用多次，就"平均"来说其偏差为零. 在科学技术中 $E(\hat{\theta})-\theta$ 称为以$\hat{\theta}$作为 θ 的估计的系统误差. 无偏估计的实际意义就是无系统误差.

结论：样本均值是总体均值的无偏估计；
样本方差是总体方差的无偏估计.

设总体X的均值（数学期望）为 $E(X)=\mu$，方差为 $D(X)=\sigma^2$

样本均值为 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ ，样本方差为 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

$$E(\overline{X})=\mu, \qquad E(S^2)=\sigma^2$$

$$\begin{aligned}
S^2&=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\\
&=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i^2-2X_i\overline{X}+\overline{X}^2\right)\\
&=\frac{1}{n-1}\sum_{i=1}^{n}X_i^2-2\overline{X}\frac{1}{n-1}\sum_{i=1}^{n}X_i+\frac{n}{n-1}\overline{X}^2\\
&=\frac{1}{n-1}\sum_{i=1}^{n}X_i^2-2\frac{n}{n-1}\overline{X}^2+\frac{n}{n-1}\overline{X}^2\\
&=\frac{1}{n-1}\sum_{i=1}^{n}X_i^2-\frac{n}{n-1}\overline{X}^2
\end{aligned}$$

样本均值是总体均值的无偏估计；样本方差是总体方差的无偏估计.

$$E(S^2)=\frac{1}{n-1}\sum_{i=1}^{n}E(X_i^2)-\frac{n}{n-1}E(\overline{X}^2)$$

$$=\frac{n}{n-1}\left[D(X_i)+(EX_i)^2\right]-\frac{n}{n-1}\left[D(\overline{X})+(E(\overline{X}))^2\right]$$

$$=\frac{n}{n-1}(\sigma^2+\mu^2)-\frac{n}{n-1}\left(\frac{1}{n}\sigma^2+\mu^2\right)$$

$$=\sigma^2$$

所以，S^2是σ^2的无偏估计量.

这也是为什么用 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$，而不用 $B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 的原因。

Q:矩估计量都是无偏估计量吗？

有前面的计算可知，对总体X的期望和方差的矩估计为

$$\begin{cases} \hat{\mu} = A_1 = \overline{X} \\ \hat{\sigma}^2 = A_2 - A_1^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \end{cases}$$

无偏估计

有偏估计

Q:矩估计量都是无偏估计量吗？

- 当只有一个参数时，我们可以通过一阶矩求解，有以下表达：

$$E(X) = f(\theta)$$

此时如果f的反函数存在，则有

$$\theta = f^{-1}(E(X))$$

此时我们定义矩估计量为随机变量

$$\hat{\theta} = f^{-1}(\overline{\mathrm{X}})$$

此时讨论无偏性则有

$$E(\hat{\theta}) = E\left(f^{-1}(\overline{\mathrm{X}})\right)$$

若函数f^{-1}满足 $E\left(f^{-1}(\overline{\mathrm{X}})\right) = f^{-1}\left(E(\overline{\mathrm{X}})\right)$，则

$$E(\hat{\theta}) = E\left(f^{-1}(\overline{\mathrm{X}})\right) = f^{-1}\left(E(\overline{\mathrm{X}})\right) = f^{-1}\left(E(\mathrm{X})\right) = f^{-1}\left(f(\theta)\right) = \theta$$

即为无偏估计。但如果不能满足，则不一定。

回顾-习题3：矩估计和最大似然估计的对比

解 (1)EX=0× θ^2+1×2θ(1-θ)+2× θ^2+3× $(1-2\theta)$ = 3-4θ

$$\theta = \frac{1}{4}(3\text{-EX}),\ \boxed{\hat{\theta}_{\mathrm{ME}}=\frac{1}{4}(3\text{-}\overline{\mathrm{X}})}$$

故θ的矩阵估计值为

$$\hat{\theta}_{\mathrm{ME}}=\frac{1}{4}(3\text{-}\bar{\mathrm{x}})=\frac{1}{4}(3\text{-}2)=\frac{1}{4}$$

(2)对给定的样本值(3,1,3,0,3,1,2,3)，似然函数为

$$L(\theta)=\prod_{i=1}^{8}P\{X_i=x_i\}=P\{X=0\}\cdot\{P(X=1)\}^2\cdot\ P\{X=2\}\cdot\{P(X=3)\}^4$$

$$=\theta^2\cdot[2\theta(1-\theta)]^2\cdot\theta^2\cdot(1-2\theta)^4=4\theta^6(1-\theta)^2(1-2\theta)^4$$

$$\mathrm{Ln}\ L=\ln 4+6\ln\theta+2\ln(1-\theta)+4\ln(1-2\theta)$$

$$\frac{d\ln L}{d\theta}=\frac{6}{\theta}-\frac{2}{1-\theta}-\frac{8}{1-2\theta}=0$$

即有$12\theta^2-14\theta+3=0$

$$得\hat{\theta}_{\mathrm{MLE}}=\frac{7-\sqrt{13}}{12}\ (\hat{\theta}=\frac{7+\sqrt{13}}{12}\ 舍去)$$

回46 回53

回顾-例题7：矩估计 vs 最大似然估计

设 X 的概率密度为

$$f(x,\alpha)=\begin{cases}(\alpha+1)x^{\alpha}, & 0<x<1\\ 0, & 其他\end{cases}$$

其中$\alpha>-1$为未知参数，$(X_1,\dots,X_n)$是取自总体 X 的样本，试求α的矩阵估计量与最大似然估计量。

解：$EX=\int_0^1 x(\alpha+1)x^{\alpha}dx=\frac{\alpha+1}{\alpha+2}$

$$\alpha=\frac{2EX-1}{1-EX},\quad \boxed{\hat{\alpha}_{ME}=\frac{2\bar{X}-1}{1-\bar{X}}}$$

似然函数为

$$L(\alpha)=\prod_{i=1}^{n}(\alpha)x_i^{\alpha}=(\alpha+1)^n(\prod_{i=1}^{n}x_i)^{\alpha}$$

$$Ln\,L=nln(\alpha+1)+\alpha\sum_{i=1}^{n}lnx_i$$

似然方程

$$\frac{dLn\,L}{d\alpha}=\frac{n}{\alpha+1}+\sum_{i=1}^{n}lnx_i=0$$

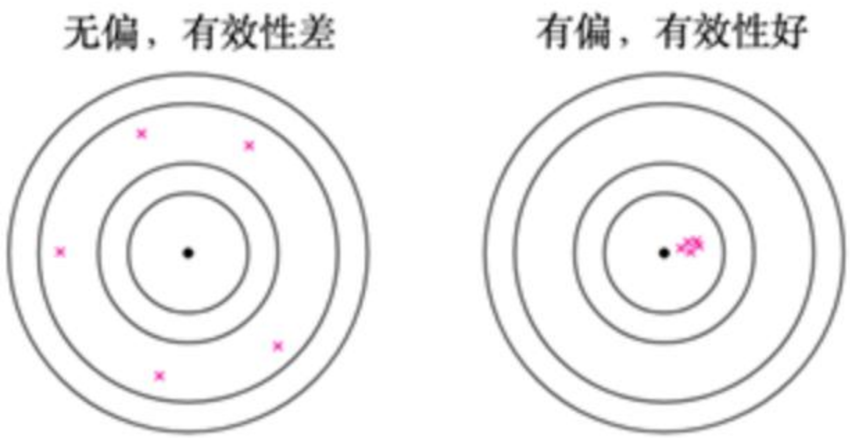
解得

$$\hat{\alpha}_{MLE}=-1-\frac{n}{\sum_{i=1}^{n}lnx_i}$$

回46 回53

Q：最大似然都是无偏估计量吗？

- 最大似然估计是否是无偏的呢？
- 是否无偏的估计一定胜过有偏的呢？



有效性

无偏估计可能有多个，如何进行选择呢？

设 $\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 与 $\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 都是 θ 的无偏估计

如果在样本容量相同的情况下，$\hat{\theta}_1$ 较 $\hat{\theta}_2$ 更密集在真值θ附近，即

$$D(\hat{\theta}_1) \le D(\hat{\theta}_2)$$

则称 $\hat{\theta}_1$ 较 $\hat{\theta}_2$**有效**.

有效性

例如 $\overline{X}$ 与 X_i 均为总体均值 μ 无偏估计，

但是 $$D(\overline{X})=\frac{\sigma^2}{n}<\sigma^2=D(X_i)$$

因此，$\overline{X}$ 较 X_i 有效.

一般时候，虽然都是样本均值，但是随着样本个数的增加，估计的方差会减小，即

$$D(\overline{X}_n)=\frac{1}{n}\sigma^2>\frac{1}{n+k}\sigma^2=D(\overline{X}_{n+k})$$

相合性（一致性）

由于统计量 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 与n有关，不妨记为 $\hat{\theta}_n$，我们自然希望 n 越大时，对 θ 的估计越精确.于是有相合性（一致性）标准。

$\hat{\theta}(X_1, X_2, \cdots, X_n)$ 是参数 θ 的估计量，如对任意的 $\theta \in \Theta$ ，

当 $n \to \infty$ 时， $\hat{\theta}_n$ 按概率收敛于θ，即对任意$\varepsilon>0$， $\lim\limits_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1$

则称 $\hat{\theta}_n$ 是θ 的相合估计量（一致估计量）。

相合性（一致性）

样本均值 $\overline{X}$ 是总体均值 μ 的相合估计

由切比雪夫定理的推论可知 $\lim\limits_{n\to\infty} P\left\{ |\overline{X}-\mu|<\varepsilon \right\}=1$

从而，样本均值 $\overline{X}$ 是总体均值 μ 的相合估计.

样本 l 阶矩是总体 l 阶矩的相合估计

$$A_l=\frac{1}{n}\sum_{i=1}^{n} X_i^l \xrightarrow{P} \mu_l=E(X^l)$$

$$h(A_1,\ldots,A_k)\xrightarrow{P} h(\mu_1,\ldots,\mu_k)$$

矩估计都是相合估计

样本方差S^2是总体方差σ^2的一致估计

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

对比：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

$$=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$$

$$A_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

$$=\frac{n}{n-1}\left(\overline{X^2}-\overline{X}^2\right)$$

当 $n\to\infty$ 时，$\overline{X^2}$ 与 $\overline{X}^2$ 分别按概率收敛于$E(X^2)$与$(EX)^2$

从而S^2按概率收敛于$E(X^2)-(EX)^2=DX=\sigma^2$

所以，S^2是σ^2的一致估计值.

总结

准则	含义	矩估计	最大似然估计
无偏性			
有效性			
相合性★			

假设有独立的正态随机变量 $X_{ij} \sim N(\mu_i, \sigma^2)$，其中 μ_i, σ^2 均未知，$1 \le i \le I, 1 \le j \le J$。然后不难得到如下MLE：

$\tilde{\mu}_i = \frac{1}{J}\sum_{j=1}^{J} X_{ij} = \bar{X}_{i\cdot}$，对任意 $i = 1, 2, \ldots, I.$

$\tilde{\sigma}^2 = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(X_{ij} - \tilde{\mu}_i\right)^2.$

NOTE: 可以证明它们是MLE，因为它们是似然方程的唯一解。

然后，固定 J，让 $I \to \infty$，得到 $\tilde{\sigma}^2 \xrightarrow{P} \frac{J-1}{J}\sigma^2$，从而不是consistent的。

我们已经讨论过了基本的参数估计方法的有偏性，且发现最大似然的有偏性估算有些困难。那么我看一些更为广泛的估计量设计及评估

例1 设总体X~N(μ,σ^2)，$(X_1,\cdots,X_n)$是取自总体X的样本，试选择适当的常数C，使$C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$为$\sigma^2$的无偏估计.

$$E(X_{i+1}X_i) = E(X_{i+1})\cdot E(X_i) = E(X)^2 = \mu^2$$

$$E(X_{i+1}^2) = E\left(X_i^2\right) = E(X^2) = \sigma^2 + \mu^2$$

$$\sigma^2 = E\left(c\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2\right) = CE\left(\sum_{i=1}^{n-1}(X_{i+1}^2 + X_i^2 - 2X_{i+1}X_i)\right)$$

$$= C\sum_{i=1}^{n-1}\left(EX_{i+1}^2 + EX_i^2 - 2EX_{i+1}X_i\right)$$

$$= C\sum_{i=1}^{n-1}\left(\mu^2 + \sigma^2 + \mu^2 + \sigma^2 - 2\mu^2\right) = 2(n-1)\sigma^2 C$$

$$C = \frac{1}{2(n-1)}$$

$\frac{1}{2(n-1)}\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2$ 为σ^2的无偏估计

这是一个我们人工编制的统计量, 但是一个无偏估计量

例2： 设总体X服从参数为λ的泊松分布，$(X_1, \cdots, X_n)$是取自总体X的样本.

(1) 试证对任意常数C, $C\overline{X}+(1-C)S^2$均是λ的无偏估计.

(2) 给出λ^2的一个无偏估计.

解(1)由于X~p(λ),EX=DX=λ,故有$E\overline{X}$=EX=λ， ES^2=DX=λ

从而对于任意C，有E(C$\overline{X}$+(1−C)S^2)=CE$\overline{X}$+(1−C)ES^2=λ

即C$\overline{X}$+(1−C)S^2均为λ的无偏估计;

（2） 由于 $E\bar{X} = \lambda$

$$E\bar{X}^2 = D\bar{X} + (E\bar{X})^2 = \frac{DX}{n} + (EX)^2 = \frac{\lambda}{n} + \lambda^2$$

取$\hat{\lambda}^2$=$\overline{X}^2-\frac{\overline{X}}{n}$, 有E$\hat{\lambda}^2$=$\lambda^2$， $\hat{\lambda}^2$为λ^2的一个无偏估计。

例3

设$(X_1,\cdots,X_n)$是取自总体X的样本，记$\mu=EX$，$\sigma^2=DX$，试证对任意常数a_i，满足$\sum_{i=1}^{n}a_i=1$时，对于$i=1,\cdots,n$, 都有 $\sum_{i=1}^{n}a_iX_i$均是μ的无偏估计，其中$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$最为有效（即$a_i=\frac{1}{n}$）。

例3

设$(X_1,\cdots,X_n)$是取自总体X的样本，记$\mu=EX$，$\sigma^2=DX$，试证对任意常数a_i，满足$\sum_{i=1}^{n}a_i=1$时，对于 $i=1,\cdots,n,$ 都有 $\sum_{i=1}^{n}a_iX_i$均是μ的无偏估计，其中$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$最为有效（即$a_i=\frac{1}{n}$）。

证明：

$$E\left(\sum_{i=1}^{n}a_iX_i\right)=\sum_{i=1}^{n}a_iEX_i=\sum_{i=1}^{n}a_iEX=\mu\sum_{i=1}^{n}a_i=\mu$$

故$\sum_{i=1}^{n}a_iX_i$均是μ的无偏估计。

$$D\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 DX_i = \sum_{i=1}^{n} a_i^2 DX = \sigma^2 \sum_{i=1}^{n} a_i^2$$

欲证$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$最为有效，只需证明当且仅当$a_i = \frac{1}{n}, i = 1,2,\cdots,n$时，$\sum_{i=1}^{n} a_i^2$达到最小。

由施瓦兹不等式

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \le \left(\sum_{i=1}^{n} x_i^2\right) \cdot \left(\sum_{i=1}^{n} y_i^2\right)$$

可得（令$x_i = a_i, y_i = 1$）

$$\sum_{i=1}^{n} a_i^2 \ge \frac{\left(\sum_{i=1}^{n} a_i\right)^2}{n} = \frac{1}{n}\left(\sum_{i=1}^{n} a_i\right) = 1$$

且等号成立当且仅当$a_i = \frac{1}{n}, i = 1,2,\cdots,n$

Remark：本题说明$\overline{X}$的优良性：对于任意总体X，$\overline{X}$是总体均值EX的最佳线性无偏估计（通常简记为BLUE）

例4

设总体X~U[0, θ]，其中$\theta > 0$为未知参数，$(X_1, \cdots, X_n)(\mathrm{n} \geq 2)$是取自总体X的样本，试证$\hat{\theta}_2 = \frac{n+1}{n}\max\limits_{1\le i\le n} X_i$ 比 $\hat{\theta}_1 = 2\overline{X}$ 有效。

例4

设总体X~U[0, θ]，其中$\theta > 0$为未知参数，$(X_1, \cdots, X_n)(\mathrm{n} \geq 2)$是取自总体X的样本，试证$\hat{\theta}_2 = \frac{n+1}{n}\max\limits_{1\leq i\leq n} X_i$ 比 $\hat{\theta}_1 = 2\overline{X}$ 有效。

分析：必须先确定$\max\limits_{1\leq i\leq n} X_i$的分布，才能计算出$\mathrm{E}\hat{\theta}_2$，$\mathrm{D}\hat{\theta}_2$，并由此验证$\mathrm{E}\hat{\theta}_2 = \theta$及$\mathrm{D}\hat{\theta}_2 < \mathrm{D}\hat{\theta}_1$

证明：先确定$X_{(n)} = \max\limits_{1\leq i\leq n} X_i$的分布。总体X的密度函数为

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta \\ 0, & \text{其他} \end{cases}$$

当$0 \le x \le \theta$时，由$X_{(n)}$的分布函数$F_{X_{(n)}}(x)$定义有

$$F_{X_{(n)}}(x) = \mathrm{P}\{X_{(n)} \le x\} = \mathrm{P}\{X_1 \le x, \cdots, X_n \le x\} = \prod_{i=1}^{n} \mathrm{P}\{X_i \le x\} = \{P(X \le x)\}^n = \left(\frac{x}{\theta}\right)^n$$

故$f_{X_{(n)}}(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n}, & 0 \le x \le \theta \\ 0, & \text{其他} \end{cases}$

最大似然估计 $\hat{\theta} =$ max(X) =$X_{(n)}$是有偏的！

再计算E$X_{(n)}$，$DX_{(n)}$

$$\mathrm{E}X_{(n)} = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+1}\theta,$$

$$\mathrm{E}{X_{(n)}}^2 = \int_0^\theta x^2 \cdot \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+2}\theta^2,$$

$$DX_{(n)} = \mathrm{E}{X_{(n)}}^2 - \left(EX_{(n)}\right)^2 = \frac{n}{(n+1)^2(n+2)}\theta^2$$

$$\mathrm{D}\hat{\theta}_2 = \left(\frac{n+1}{n}\right)^2 \mathrm{D}X_{(n)} = \frac{\theta^2}{n(n+2)}$$

乘$\frac{n+1}{n}$就无偏了

$$\mathrm{E}\hat{\theta}_2 = \frac{n+1}{n}\mathrm{E}X_{(n)} = \theta$$

$$\mathrm{E}\hat{\theta}_1=2E\overline{X}=2EX=2\cdot\frac{\theta}{2}=\theta,$$

$$\mathrm{D}\hat{\theta}_1=4D\overline{X}=4\cdot\frac{DX}{n}=4\cdot\frac{\theta^2}{12n}=\frac{\theta^2}{3n},$$

显然$\mathrm{E}\hat{\theta}_2=\mathrm{E}\hat{\theta}_1=\theta$，且$\mathrm{D}\hat{\theta}_2=\frac{\theta^2}{n(n+2)}<\frac{\theta^2}{3n}=\mathrm{D}\hat{\theta}_1$，

因此 $\hat{\theta}_2$比$\hat{\theta}_1$有效。

Remark：在点估计中，经常对利用各种点估计方法得到的估计量进行比较，有时还进行改进，从而得到未知参数的较优估计。

例如：本题中，$\max\limits_{1\le i\le n}X_i$与$2\overline{X}$恰为$\theta$的最大似然估计与矩估计，虽然$\max\limits_{1\le i\le n}X_i$是有偏的，但改进后的估计$\frac{n+1}{n}\max\limits_{1\le i\le n}X_i$不仅是无偏的，并且比$2\overline{X}$更有效。

主观题 10分

设置

设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}\theta, & 0<x<1\\ 1-\theta, & 1\le x<2\\ 0, & \text{其他}\end{cases}$$

其中 $0<\theta<1$ 为未知参数，$(X_1,\cdots,X_n)$ 是取自总体 X 的样本，试求 θ 的矩估计量与最大似然估计量

作答

主观题 10分

设置

设总体X的密度函数为

$$f(x;\theta)=\begin{cases}\frac{1}{2\theta}, & 0<x<\theta\\ \frac{1}{2(1-\theta)}, & \theta\leqslant x<1\\ 0, & \text{其他}\}\end{cases}$$

其中0<θ<1为未知参数，$(X_1,\cdots,X_n)$是取自总体X的样本

(1) 求θ的矩估计量$\hat{\theta}_{ME}$；

(2) 判断$4\overline{X}^2$是否为θ^2的无偏估计量.

作答

小结论

我们用各种估计方法，除了计算估计值外，更重要的是构造一个统计量

构造的统计量（包括矩估计、最大似然）的统计量未必满足最优条件（无偏、有效性、收敛性）

什么是真正意义的最佳统计量，实际上在现有体系中是“无解”的。即使引入Bayes分析也一样，需要一模型一议

不管哪一种估计方法，哪一种统计量，随着抽样数量n上升，其期望都将逼近于对应的参数 不对！！！

此时，n越大，方差一定越小；期望的有效性一定越好，无偏性就越重要。