概率论与数理统计 B　第十周作业　4 月 21 日 周二

PB18151866 龚小航

7.19. (2016 年研究生入学考试试题) 设总体 X 的概率密度为

$$f(x,\theta)=\begin{cases}\dfrac{3x^2}{\theta^3}, & 0<x<\theta\\ 0, & 其他\end{cases}$$

其中$\theta\in(0,\infty)$为未知参数, X_1,X_2,X_3为总体 X 的简单随机抽样, 令 $T=\max(X_1,X_2,X_3)$.

(1) 求 T 的概率密度;

(2) 确定 a, 使得 aT 为θ的无偏估计.

解: (1) 可以先求分布函数F_T:

$$F_T(T\le t)=\prod_{i=1}^{3}P(X_i\le t)=\prod_{i=1}^{3}F_X(t)=\left(\frac{t}{\theta}\right)^9$$

$$f_T(t,\theta)=\begin{cases}\dfrac{9t^8}{\theta^9}, & 0<t<\theta\\ 0, & 其他\end{cases}$$

(2) 样本均值是总体分布的无偏估计, 故有:

$$E(T)=\int_{-\infty}^{\infty}tf_T(t)\,dt=\int_0^{\infty}t\frac{9t^8}{\theta^9}\,dt=\frac{9}{10}\theta$$

$$\Rightarrow\quad \hat{\theta}=\frac{10}{9}\overline{T}\qquad\Rightarrow a=\frac{10}{9}$$

7.20 设 $X_1,\dots,X_n$ 是来自总体 X 的一个简单随机样本, $E(X)=\mu$, $Var(X)=\sigma^2$.

(1) 确定常数 c 使得 $c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计.

(2) 记 $\overline{X},S^2$ 分别是样本均值和样本方差. 确定常数 c 使 $\overline{X}^2-cS^2$ 是 μ^2 的无偏估计.

解: (1) 样本均值是总体分布的无偏估计, 故有:

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

再对这三项进行计算后带入:

$$E(X)=\mu,\quad Var(X)=E(X^2)-E^2(X)=\sigma^2\quad\Rightarrow\quad E(X^2)=\sigma^2+\mu^2;$$

$$E(X_iX_{i+1})=E(X_i)E(X_{i+1});$$

因此:

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

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

(2) 先写出$\overline{X},S^2$的表达式, 再用样本均值计算无偏估计: 其中$E(X_iX_j)=E(X_i)E(X_j)$

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

$$E\left(\overline{X}^2\right)=E\left(\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right)=\frac{1}{n^2}E\left(\sum_{i=1}^{n}X_i^2+2\sum_{i\ne j}X_iX_j\right)=\frac{1}{n^2}\sum_{i=1}^{n}E(X_i^2)+\frac{2}{n^2}\sum_{i\ne j}E(X_iX_j)$$

$$=\frac{1}{n^2}nE(X_i^2)+\frac{2}{n^2}\frac{n(n-1)}{2}E(X_i)E(X_j)=\frac{1}{n^2}n(\sigma^2+\mu^2)+\frac{2}{n^2}\frac{n(n-1)}{2}\mu^2\quad=\quad\frac{\sigma^2}{n}+\mu^2$$

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

其中, 对最后一项变形计算:

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

$$E(S^2)=\frac{1}{n-1}\left(\sum_{i=1}^{n}E\left(X_i^2\right)-E\left(n\sum_{i=1}^{n}\overline{X}^2\right)\right)=\frac{1}{n-1}*n(\sigma^2+\mu^2)-\frac{1}{n-1}*n\left(\frac{\sigma^2}{n}+\mu^2\right)=\frac{n\sigma^2}{n-1}\left(1-\frac{1}{n}\right)$$

最后再将这两项带入:

$$E\left(\overline{X}^2-cS^2\right)=E\left(\overline{X}^2\right)-cE(S^2)=\frac{\sigma^2}{n}+\mu^2-c\frac{n\sigma^2}{n-1}\left(1-\frac{1}{n}\right)\equiv\mu^2$$

$$\Rightarrow\quad c=\frac{1}{n}$$

7.21. 设从均值为μ, 方差为σ^2的总体中, 分别抽取容量为n_1, n_2的两个独立样本. 设 X_1, X_2分别是两样本的均值. 试证明对于任意常数a, $Y = a\overline{X_1} + (1-a)\overline{X_2}$是 μ 的无偏估计, 并确定常数$a$使$Y$的方差达到最小.

解：直接写出Y的期望即可：

$$E(Y) = E\left(a\overline{X_1} + (1-a)\overline{X_2}\right) = aE\left(\overline{X_1}\right) + (1-a)E\left(\overline{X_2}\right) = aE\left(\frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i}\right) + (1-a)E\left(\frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i}\right)$$

$$= \frac{a}{n_1}\sum_{i=1}^{n_1} E(X_{1i}) + \frac{1-a}{n_2}\sum_{i=1}^{n_2} E(X_{2i}) = \frac{a}{n_1} * n_1 \mu + \frac{1-a}{n_2} * n_2 \mu = \mu \equiv \mu$$

因此显然$Y = a\overline{X_1} + (1-a)\overline{X_2}$是 μ 的无偏估计。

再求Y的方差：

$$Var(Y) = Var\left(a\overline{X_1} + (1-a)\overline{X_2}\right) = a^2 Var\left(\overline{X_1}\right) + (1-a)^2 Var\left(\overline{X_2}\right)$$

$$= a^2 Var\left(\frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i}\right) + (1-a)^2 Var\left(\frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i}\right) = \frac{a^2}{n_1^2}\sum_{i=1}^{n_1} Var(X_{1i}) + \frac{(1-a)^2}{n_2^2}\sum_{i=1}^{n_2} Var(X_{2i})$$

$$= \frac{a^2}{n_1^2} * n_1\sigma^2 + \frac{(1-a)^2}{n_2^2} * n_2\sigma^2 = \sigma^2\left(\frac{a^2}{n_1} + \frac{(1-a)^2}{n_2}\right) = \sigma^2\left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right)a^2 - \frac{2}{n_2}a + \frac{1}{n_2}\right)$$

$$= \frac{n_1+n_2}{n_1 n_2}\sigma^2\left(a^2 - \frac{2n_1}{n_1+n_2}a + \frac{n_1}{n_1+n_2}\right) = \frac{n_1+n_2}{n_1 n_2}\sigma^2\left(\left(a - \frac{n_1}{n_1+n_2}\right)^2 + k\right)$$

这二次函数显然极小值点在$a = \frac{n_1}{n_1+n_2}$时取到，因此$a = \frac{n_1}{n_1+n_2}$

7.27. 一袋中有N 个均匀硬币, 其中θ个是普通的硬币, 其余 $N-\theta$ 个两面都是正面. 现从袋中随机摸出一个把它连掷两次, 记下结果, 但是不看它属于哪种硬币, 又把它放回袋中, 如此重复n次. 如果掷出 0,1,2 次正面的次数分别是n_0, n_1, n_2 次 $(n_0 + n_1 + n_2 = n)$, 试分别用矩估计法和极大似然法这两种方法估计袋中普通硬币数θ.

解：先从总体描述写出事件X：抛掷两次硬币，硬币朝上次数。 再求出X的期望：

① 0 次正面：（2 次反面）必须选中普通硬币，对应的概率为：

$$P(X=0) = \frac{\theta}{N} * \frac{1}{2} * \frac{1}{2} = \frac{\theta}{4N}$$

② 1 次正面：（一正一反）由于有反面存在，必须选中普通硬币，对应的概率为：

$$P(X=1) = \frac{\theta}{N} * \left(\frac{1}{2} * \frac{1}{2} + \frac{1}{2} * \frac{1}{2}\right) = \frac{\theta}{2N}$$

③ 2 次正面：（0 次反面）必须选中普通硬币，对应的概率为：

$$P(X=2) = 1 - P(X=1) - P(X=0) = 1 - \frac{3\theta}{4N}$$

由此可以写出X的期望：

$$E(X) = 0 * \frac{\theta}{4N} + 1 * \frac{\theta}{2N} + 2 * \left(1 - \frac{3\theta}{4N}\right) = 2 - \frac{\theta}{N}$$

然后从样本的角度出发：

矩估计：利用样本矩估计总体矩θ：令样本均值为a

$$a = \frac{0 * n_0 + 1 * n_1 + 2 * n_2}{n} = \frac{n_1 + 2n_2}{n} = E(X) = 2 - \frac{\theta}{N}$$

$$\Longrightarrow \quad \hat{\theta} = N\left(2 - \frac{n_1 + 2n_2}{n}\right) = \frac{N}{n}(2n - n_1 - 2n_2) = \frac{N}{n}(2n_0 + n_1)$$

极大似然估计：由样本写出似然函数：

$$L(\theta) = \prod_{i=0}^{2} P(x,\theta) = \left(\frac{\theta}{4N}\right)^{n_0}\left(\frac{\theta}{2N}\right)^{n_1}\left(1 - \frac{3\theta}{4N}\right)^{n_2} \quad 换元，令t = \frac{\theta}{4N}：$$

$$L(t) = t^{n_0}(2t)^{n_1}(1-3t)^{n_2} = 2^{n_1}t^{n_0+n_1}(1-3t)^{n_2} = 2^{n_1}t^{n-n_2}(1-3t)^{n_2}$$

$$\ln L = n_1 \ln 2 + (n - n_2)\ln t + n_2 \ln(1-3t)$$

再对其求导寻找极大值：

$$\frac{\partial \ln L}{\partial t} = \frac{n - n_2}{t} - \frac{3n_2}{1-3t} = 0 \quad \Longrightarrow \quad t = \frac{n - n_2}{3n}$$

因此参数θ的极大似然估计为：

$$\hat{\theta} = 4Nt = 4N\frac{n - n_2}{3n}$$

7.42. 设 $X_1,\dots,X_n$ 为来自均匀分布总体$U(0,\theta)$的一个简单随机样本，求θ的极大似然估计$\hat{\theta}_n$

并且证明$\hat{\theta}_n$是θ的相合估计。

解：似然函数可以直接写出：

$$L(\theta)=\prod_{i=1}^{n}P(x_i,\theta)=\begin{cases}\dfrac{1}{\theta^n}\ , & 0\le X_i\le\theta\\ 0\ , & \text{其他}\end{cases}$$

似然函数是单调递减的，所以θ应该取能取到的最小值。而限制条件为$0\le X_i\le\theta$规定了θ的下界。

重排$X_1,\dots,X_n$，使$X^{(1)}\le X^{(2)}\le,\dots,\le X^{(n)}$，所以$\theta$的下界即为$X^{(n)}$

$$\Longrightarrow\quad \hat{\theta}_n=X^{(n)}$$

再证明$\hat{\theta}_n$是θ的相合估计。按定义，需要证明下式：

$$\lim_{n\to\infty}P_\theta(\left|\hat{\theta}_n-\theta\right|\ge\varepsilon)=0$$

先写出从总体看$X^{(n)}$的分布函数与概率密度函数：

$$F_{X^{(n)}}(x)=P\left(X^{(n)}\le x\right)=\prod_{i=1}^{n}P(X_i\le x)=\left(\frac{x}{\theta}\right)^n\quad ;\quad f_{X^{(n)}}(x)=\frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1}\quad 0\le x\le\theta$$

再写出极大似然估计的均值：

$$E\left(\hat{\theta}_n\right)=E\left(X^{(n)}\right)=\int_{-\infty}^{\infty}xf_{X^{(n)}}(x)\,dx=\int_0^\theta x\frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1}dx=\frac{n}{n+1}\theta$$

从上述表达式可知，$\lim\limits_{n\to\infty}E\left(\hat{\theta}_n\right)=\theta$ 利用契比雪夫不等式，有：

$$\lim_{n\to\infty}P_\theta(\left|\hat{\theta}_n-\theta\right|\ge\varepsilon)=\lim_{n\to\infty}P_\theta(\left|\hat{\theta}_n-E\left(\hat{\theta}_n\right)\right|\ge\varepsilon)\le\frac{Var(\hat{\theta}_n)}{\varepsilon^2}$$

因此还需再求出$Var(\hat{\theta}_n)$：

$$Var\left(\hat{\theta}_n\right)=E\left(\hat{\theta}_n^2\right)-E^2\left(\hat{\theta}_n\right)=\int_{-\infty}^{\infty}x^2f_{X^{(n)}}(x)\,dx-\left(\frac{n}{n+1}\theta\right)^2=\left(\frac{n}{n+2}-\frac{n^2}{(n+1)^2}\right)\theta^2$$

带入不等式中，利用夹逼定理，可得：

$$0\le\lim_{n\to\infty}P_\theta\left(\left|\hat{\theta}_n-\theta\right|\ge\varepsilon\right)\le\frac{Var\left(\hat{\theta}_n\right)}{\varepsilon^2}=\frac{1}{\varepsilon^2}\lim_{n\to\infty}\left(\frac{n}{n+2}-\frac{n^2}{(n+1)^2}\right)\theta^2=0$$

此即：

$$\lim_{n\to\infty}P_\theta(\left|\hat{\theta}_n-\theta\right|\ge\varepsilon)=0$$

因此$\hat{\theta}_n$是θ的相合估计