第 七 章

1.设 X 的概率密度为 $f(x)=\begin{cases}\dfrac{x}{\theta^2}e^{-x^2/(2\theta^2)}, & x>0\\ 0, & x\le 0\end{cases}$，求 θ 的矩估计量和极大似然估计量.

解：矩估计

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^{+\infty}x\frac{x}{\theta^2}e^{-x^2/(2\theta^2)}dx=-\int_0^{+\infty}xde^{-x^2/(2\theta^2)}$$

$$=\int_0^{+\infty}e^{-x^2/(2\theta^2)}dx=\theta\int_0^{+\infty}e^{-\frac{1}{2}\left(\frac{x}{\theta}\right)^2}d\frac{x}{\theta}=\frac{\theta\sqrt{2\pi}}{2}$$

令 $E(X)=\dfrac{\theta\sqrt{2\pi}}{2}=\overline{X}$，得 θ 的矩估计量 $\hat{\theta}=\sqrt{\dfrac{2}{\pi}}\overline{X}$

极大似然估计：

设样本观测值为 $x_1,x_2,\cdots,x_n$

似然函数 $L(\theta)=\begin{cases}\displaystyle\prod_{i=1}^{n}\frac{x_i}{\theta^2}e^{-x_i^2/(2\theta^2)}=\theta^{-2n}e^{-\frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2}\prod_{i=1}^{n}x_i, & x_i>0(i=1,\ldots,n)\\ 0, & \text{其它}\end{cases}$

当 $x_i>0(i=1,\ldots,n)$ 时，$\ln L=-2n\ln\theta-\dfrac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}\ln x_i$

令 $\dfrac{d\ln L}{d\theta}=\dfrac{-2n}{\theta}+\dfrac{1}{\theta^3}\sum_{i=1}^{n}x_i^2=0$

得 θ 的极大似然估计值 $\hat{\theta}=\sqrt{\dfrac{1}{2n}\sum_{i=1}^{n}x_i^2}$，极大似然估计量 $\hat{\theta}=\sqrt{\dfrac{1}{2n}\sum_{i=1}^{n}X_i^2}$

2.设总体 X 服从几何分布：$P\{X=k\}=(1-p)^{k-1}p\quad(0<p<1),\quad k=1,2,3,\cdots$,求 p 的极大似然估计量.

解：总体分布律写成 $P\{X=x\}=(1-p)^{x-1}p\quad(0<p<1),\quad x=1,2,3,\cdots$

似然函数

$$L(p)=\prod_{i=1}^{n}(1-p)^{x_i-1}p=p^n(1-p)^{\sum_{i=1}^{n}x_i-n},\quad x_i=1,2,3,\ldots(i=1,\ldots,n)$$

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

令 $\dfrac{d\ln L}{dp}=\dfrac{n}{p}-\dfrac{\sum_{i=1}^{n}x_i-n}{1-p}=0$，得 p 的极大似然估计值 $\hat{p}=\dfrac{n}{\sum_{i=1}^{n}x_i}=\dfrac{1}{\overline{x}}$

极大似然估计量 $\hat{p}=\dfrac{n}{\sum_{i=1}^{n}X_i}=\dfrac{1}{\overline{X}}$

3. 设总体 X 的分布律为：

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

其中 $0<\theta<1/2$ 为未知参数，利用如下样本值：3， 1， 0， 3， 2 求 θ 的矩估计值和最大似然估计值.

解：矩估计：

$$E(X)=2\theta(1-\theta)+2\theta^2+3(1-2\theta)=3-4\theta$$

令 $E(X)=3-4\theta=\overline{X}$ 得 θ 的矩估计量 $\hat{\theta}=\dfrac{3-\overline{X}}{4}$，代入样本观测值得 $\overline{x}=1.8$，从而

矩估计值 $\hat{\theta}=0.3$

极大似然估计　似然函数为

$$\begin{aligned}L(\theta)&=P\{X_1=3,X_2=1,X_3=0,X_4=3,X_5=2\}\\&=P\{X_1=3\}P\{X_2=1\}\ldots P\{X_5=2\}\\&=2\theta^5(1-\theta)(1-2\theta)^2\end{aligned}$$

$$\ln L=\ln 2+5\ln\theta+\ln(1-\theta)+2\ln(1-2\theta)$$

令 $\dfrac{d\ln L}{d\theta}=\dfrac{5}{\theta}-\dfrac{1}{1-\theta}-\dfrac{4}{1-2\theta}=0\Rightarrow 16\theta^2-20\theta+5=0$，

注意 $0<\theta<1/2$，得 θ 的极大似然估计值 $\hat{\theta}=\dfrac{5-\sqrt{5}}{8}\approx 0.345$

4.设总体 $X\sim N(\mu,\sigma^2)$， $X_1,X_2,\cdots,X_n$ 为其样本，试求常数C使 $C\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计量。

解：$E(C\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2)=C\sum\limits_{i=1}^{n-1}[E(X_{i+1}^2)-2E(X_{i+1}X_i)+E(X_i^2)]$

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

$$=C[(n-1)\sigma^2+(n-1)\mu^2-2(n-1)\mu^2+(n-1)\sigma^2+(n-1)\mu^2]$$

$$=C\cdot 2(n-1)\sigma^2=\sigma^2\Rightarrow C=\frac{1}{2(n-1)}$$

5. 设总体 $X\sim N(\mu,1)$，X_1,X_2 为其样本，问：估计量 $\hat{\mu}_1=\dfrac{2}{3}X_1+\dfrac{1}{3}X_2$, $\hat{\mu}_2=\dfrac{1}{2}X_1+\dfrac{1}{2}X_2$, $\hat{\mu}_3=\dfrac{1}{3}X_1+\dfrac{1}{2}X_2$ 中，哪一个是 μ 的较有效估计量？

解：$E(\hat{\mu}_1)=E(\dfrac{2}{3}X_1+\dfrac{1}{3}X_2)=\dfrac{2}{3}E(X_1)+\dfrac{1}{3}E(X_2)=\dfrac{2}{3}E(X)+\dfrac{1}{3}E(X)=\mu$

$$E(\hat{\mu}_2)=E(\frac{1}{2}X_1+\frac{1}{2}X_2)=\frac{1}{2}E(X_1)+\frac{1}{2}E(X_2)=\frac{1}{2}E(X)+\frac{1}{2}E(X)=\mu$$

$$E(\hat{\mu}_3)=E(\frac{1}{3}X_1+\frac{1}{2}X_2)=\frac{1}{3}E(X_1)+\frac{1}{2}E(X_2)=\frac{1}{3}E(X)+\frac{1}{2}E(X)=\frac{5}{6}\mu$$

故只有 $\hat{\mu}_1,\hat{\mu}_2$ 是 μ 的无偏估计量，下面对它们考察有效性

$$D(\hat{\mu}_1)=D(\frac{2}{3}X_1+\frac{1}{3}X_2)=\frac{4}{9}D(X_1)+\frac{1}{9}D(X_2)=\frac{5}{9}$$

$$D(\hat{\mu}_2)=D(\frac{1}{2}X_1+\frac{1}{2}X_2)=\frac{1}{4}D(X_1)+\frac{1}{4}D(X_2)=\frac{1}{2}$$

$\hat{\mu}_2$ 比较有效。

6.设总体 $X\sim B(1,p)$, $X_1,X_2,\cdots X_n$ 为其样本,验证统计量 $T=\frac{1}{n}(X_1^2+X_2^2+\cdots+X_n^2)$ 是参数 p 的相合估计量.

证明：注意到 X 与 X^2 具有相同分布，$X\sim B(1,p)\Rightarrow E(X^2)=p$

由 $X_1,X_2,\cdots X_n$ 的独立同分布性可知 $X_1^2,X_2^2,\cdots X_n^2$ 仍然独立同分布，满足独立同分布大数定律的条件，而 $E(T)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\frac{1}{n}\sum_{i=1}^{n}E(X^2)=p$ 故

$$T=\frac{1}{n}(X_1^2+X_2^2+\cdots+X_n^2)\xrightarrow{p}p(n\to\infty)$$

即统计量 $T=\frac{1}{n}(X_1^2+X_2^2+\cdots+X_n^2)$ 是参数 p 的相合估计量

注：也可用切比雪夫不等式直接证明

$$E(T)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\frac{1}{n}\sum_{i=1}^{n}E(X^2)=p,$$

$$D(T)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i^2)=\frac{1}{n^2}\sum_{i=1}^{n}D(X^2)=\frac{1}{n^2}\sum_{i=1}^{n}(p-p^2)=\frac{p-p^2}{n}$$

$$P\{|T-p|\ge\varepsilon\}\le\frac{D(T)}{\varepsilon^2}=\frac{p-p^2}{n\varepsilon^2}\to 0$$

7. 设某种清漆的干燥时间（小时）服从正态分布 $N(\mu,\sigma^2)$，现有一组样本观测值：

6.0, 5.7, 5.8, 6.5, 7.0, 6.3, 5.6, 6.1, 5.0

求 μ 的置信度为 0.95 的置信区间。（1）已知 $\sigma=0.6$；（2）σ 未知。

解： 1)需估计μ，σ^2已知，故选枢轴变量：$U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

令 $P\{|U|\le u_{\alpha/2}\}=1-\alpha$

解得 μ 的置信区间 $[\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\ \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}]$

$\alpha=0.05, n=9, \bar{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=6,\ u_{0.025}=1.96$,得$\mu$的置信度为0.95的置信区间为:

$$[6-\frac{0.6}{3}\times1.96,\ 6+\frac{0.6}{3}\times1.96]=[5.608,\ 6.392]$$

2) 需估计μ，σ^2未知，故选枢轴变量：$T=\frac{\overline{X}-\mu}{S/\sqrt{n-1}}\sim t(n-1)$

令 $P\{|T|\le t_{\alpha/2}(n-1)\}=1-\alpha$

解得 μ 的置信区间 $[\overline{X}-\frac{S}{\sqrt{n-1}}t_{\alpha/2}(n-1),\ \overline{X}+\frac{S}{\sqrt{n-1}}t_{\alpha/2}(n-1)]$

$\alpha=0.05, n=9, \bar{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=6, s=\sqrt{\frac{1}{9}\sum_{i=1}^{9}(x_i-\bar{x})^2}\approx0.5416,$

$t_{0.025}(8)=2.306$,得μ的置信度为0.95的置信区间为:

$$[6-\frac{0.5416}{2\sqrt{2}}\times2.306,\ 6+\frac{0.5416}{2\sqrt{2}}\times2.306]=[5.558,\ 6.442]$$

8.某商店一种产品的月销售量服从正态分布 $N(\mu,\sigma^2)$，随机抽取 7 个月的销售量观察：64, 57, 49, 81, 76, 70, 59, 求 σ^2 的置信度为 0.9 的置信区间.

解：设 X 为产品的月销售量，服从正态分布 $N(\mu,\sigma^2)$，其样本为 $X_1,\cdots,X_7$，

要估计总体方差，均值未知，选取枢轴变量 $\chi^2=\dfrac{(n-1)\ S^2}{\sigma^2}\sim\chi^2(n-1)$

令 $$P\{\chi^2{}_{1-\alpha/2}(n-1)\le\chi^2\le\chi^2{}_{\alpha/2}(n-1)\}=1-\alpha$$

解得 σ^2 的置信区间 $\left[\dfrac{(n-1)S^2}{\chi^2{}_{\alpha/2}(n-1)},\ \dfrac{(n-1)S^2}{\chi^2{}_{1-\alpha/2}(n-1)}\right]$

由样本观测值算得 $s^2=$ 108.4082， $\alpha=0.1$， $\chi^2{}_{0.05}(6)=12.592,\chi^2{}_{0.95}(6)=1.635$，

故置信区间为 $\left[\dfrac{(n-1)S^2}{\chi^2{}_{\alpha/2}(n-1)},\ \dfrac{(n-1)S^2}{\chi^2{}_{1-\alpha/2}(n-1)}\right]=[\,60.27,\ \ 464.13]$

9. 对方差 σ^2 为已知的正态总体，问：需取容量 n 为多大的样本才能使总体均值的置信度 $1-\alpha$ 的置信区间长度不大于 L？

解：由于方差已知，选取枢轴变量 $U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

令 $$P\{|U|\le u_{\alpha/2}\}=1-\alpha$$

则总体均值的置信度为 1－α的置信区间为： $\left[\overline{X}-\dfrac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}},\ \overline{X}+\dfrac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right]$

区间长度为：$2\dfrac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\le L\ \Rightarrow n\ge(2u_{\frac{\alpha}{2}}\dfrac{\sigma}{L})^2=\dfrac{4\sigma^2}{L^2}(u_{\frac{\alpha}{2}})^2$

10.设某地区男、女身高 X、Y 均服从正态分布且方差相等，随机抽取成人男、女各 100 名，测量并计算得男子身高 $\bar{x}=1.71m,s_1=0.035m$，女子身高 $\bar{y}=1.67m,s_2=0.038m$. 求男、女平均身高之差的置信度 0.95 的置信区间.

11.测量值分别算得 $s_1^2=0.5419,s_2^2=0.6065$，设总体均为正态分布，求方差比 σ_1^2/σ_2^2 的置信度 95%的置信区间。

12.从某型号的一批电子管中抽出容量为 10 的样本做寿命试验，算得 $S=45$（小时），设整批电子管的寿命服从正态分布，试求这批电子管寿命标准差的单侧置信上限（置信度为 0.95）.