习题 6.3

1. 设总体概率函数如下，$X_1, \cdots, X_n$ 是样本，试求未知参数的最大似然估计。

（1） $p(x;\theta)=\sqrt{\theta}x^{\sqrt{\theta}-1}$，$0<x<1$，$\theta>0$；

（2）$p(x;\theta)=\theta c^{\theta}x^{-(\theta+1)}$，$x>c$，$c>0$ 已知，$\theta>1$。

解：（1）因 $L(\theta)=\prod\limits_{i=1}^{n}\sqrt{\theta}x_i^{\sqrt{\theta}-1}\mathrm{I}_{0<x_i<1}=\theta^{\frac{n}{2}}(x_1x_2\cdots x_n)^{\sqrt{\theta}-1}\mathrm{I}_{0<x_1,\,x_2,\,\cdots,\,x_n<1}$，

当 $0<x_1, x_2, \cdots, x_n<1$ 时，$\ln L(\theta)=\dfrac{n}{2}\ln\theta+(\sqrt{\theta}-1)\ln(x_1x_2\cdots x_n)$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=\dfrac{n}{2\theta}+\dfrac{1}{2\sqrt{\theta}}\ln(x_1x_2\cdots x_n)=0$，得 $\sqrt{\theta}=-\dfrac{n}{\ln(x_1x_2\cdots x_n)}$，即 $\theta=\left[\dfrac{n}{\ln(x_1x_2\cdots x_n)}\right]^2$，

故 θ 的最大似然估计 $\hat{\theta}=\left[\dfrac{n}{\ln(X_1X_2\cdots X_n)}\right]^2$；

（2）因 $L(\theta)=\prod\limits_{i=1}^{n}\theta c^{\theta}x_i^{-(\theta+1)}\mathrm{I}_{x_i>c}=\theta^n c^{n\theta}(x_1x_2\cdots x_n)^{-(\theta+1)}\mathrm{I}_{x_1,\,x_2,\,\cdots,\,x_n>c}$，

当 $x_1, x_2, \cdots, x_n>c$ 时，$\ln L(\theta)=n\ln\theta+n\theta\ln c-(\theta+1)\ln(x_1x_2\cdots x_n)$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=\dfrac{n}{\theta}+n\ln c-\ln(x_1x_2\cdots x_n)=0$，得 $\theta=\dfrac{n}{\ln(x_1x_2\cdots x_n)-n\ln c}$，

故 θ 的最大似然估计 $\hat{\theta}=\dfrac{n}{\ln(X_1X_2\cdots X_n)-n\ln c}$。

2. 设总体概率函数如下，$X_1, \cdots, X_n$ 是样本，试求未知参数的最大似然估计。

（1）$p(x;\theta)=c\theta^{c}x^{-(c+1)}$，$x>\theta$，$\theta>0$，$c>0$ 已知；

（2） $p(x;\theta,\mu)=\dfrac{1}{\theta}\mathrm{e}^{-\frac{x-\mu}{\theta}}$，$x>\mu$，$\theta>0$；

（3）$p(x;\theta)=(k\theta)^{-1}$，$\theta<x<(k+1)\theta$，$\theta>0$。

解：（1）因 $L(\theta)=\prod\limits_{i=1}^{n}c\theta^{c}x_i^{-(c+1)}\mathrm{I}_{x_i>\theta}=c^n\theta^{nc}(x_1x_2\cdots x_n)^{-(c+1)}\mathrm{I}_{x_1,x_2,\cdots,x_n>\theta}$，

显然 θ 越大，θ^{nc} 越大，但只有 $x_1, x_2, \cdots, x_n>\theta$ 时，才有 $L(\theta)>0$，
即 $\theta=\min\{x_1, x_2, \ldots, x_n\}$ 时，$L(\theta)$ 达到最大，

故 θ 的最大似然估计 $\hat{\theta}=X_{(1)}=\min\{X_1, X_2, \cdots, X_n\}$；

（2）因 $L(\theta,\mu)=\prod\limits_{i=1}^{n}\dfrac{1}{\theta}\mathrm{e}^{-\frac{x_i-\mu}{\theta}}\mathrm{I}_{x_i>\mu}=\dfrac{1}{\theta^n}\mathrm{e}^{-\frac{1}{\theta}\left(\sum\limits_{i=1}^{n}x_i-n\mu\right)}\mathrm{I}_{x_1,x_2,\cdots,x_n>\mu}$，

当 $x_1, x_2, \cdots, x_n>\mu$ 时，$\ln L(\theta,\mu)=-n\ln\theta-\dfrac{1}{\theta}\left(\sum\limits_{i=1}^{n}x_i-n\mu\right)$，

令 $\dfrac{d\ln L(\theta,\mu)}{d\theta}=-\dfrac{n}{\theta}+\dfrac{1}{\theta^2}\left(\sum_{i=1}^n x_i-n\mu\right)=0$，解得 $\theta=\dfrac{1}{n}\left(\sum_{i=1}^n x_i-n\mu\right)=\bar{x}-\mu$，

且显然 μ 越大，$\mathrm{e}^{-\frac{1}{\theta}\left(\sum_{i=1}^n x_i-n\mu\right)}$ 越大，但只有 $x_1, x_2, \cdots, x_n>\mu$ 时，才有 $L(\theta,\mu)>0$，

即 $\mu=\min\{x_1, x_2, \ldots, x_n\}$ 时，$L(\theta,\mu)$ 才能达到最大，

故 μ 的最大似然估计 $\hat{\mu}=X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$，$\theta$ 的最大似然估计 $\hat{\theta}=\bar{X}-\hat{\mu}=\bar{X}-X_{(1)}$；

（3）因 $L(\theta)=\prod_{i=1}^n (k\theta)^{-1}\mathrm{I}_{\theta<x_i<(k+1)\theta}=(k\theta)^{-n}\mathrm{I}_{\theta<x_1,x_2,\cdots,x_n<(k+1)\theta}$，

显然 θ 越小，$(k\theta)^{-n}$ 越大，但只有 $\theta<x_1, x_2, \cdots, x_n<(k+1)\theta$ 时，才有 $L(\theta)>0$，

即 $\theta=\dfrac{1}{k+1}\max\{x_1,x_2,\cdots,x_n\}$ 时，$L(\theta)$ 达到最大，

故 θ 的最大似然估计为 $\hat{\theta}=\dfrac{X_{(n)}}{k+1}=\dfrac{1}{k+1}\max\{X_1,X_2,\cdots,X_n\}$。

3．设总体概率函数如下，$X_1, \cdots, X_n$ 是样本，试求未知参数的最大似然估计。

（1）$p(x;\theta)=\dfrac{1}{2\theta}\mathrm{e}^{-|x|/\theta}$，$\theta>0$；

（2）$p(x;\theta)=1$，$\theta-1/2<x<\theta+1/2$；

（3）$p(x;\theta_1,\theta_2)=\dfrac{1}{\theta_2-\theta_1}$，$\theta_1<x<\theta_2$。

解：（1）因 $L(\theta)=\prod_{i=1}^n\dfrac{1}{2\theta}\mathrm{e}^{-|x_i|/\theta}=\dfrac{1}{2^n\theta^n}\mathrm{e}^{-\frac{1}{\theta}\sum_{i=1}^n|x_i|}$，有 $\ln L(\theta)=-n\ln 2-n\ln\theta-\dfrac{1}{\theta}\sum_{i=1}^n|x_i|$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=-n\cdot\dfrac{1}{\theta}+\dfrac{1}{\theta^2}\sum_{i=1}^n|x_i|$，得 $\theta=\dfrac{1}{n}\sum_{i=1}^n|x_i|$，

故 θ 的最大似然估计 $\hat{\theta}=\dfrac{1}{n}\sum_{i=1}^n|X_i|$；

（2）因 $L(\theta)=\prod_{i=1}^n\mathrm{I}_{\theta-1/2<x_i<\theta+1/2}=\mathrm{I}_{\theta-1/2<x_1,x_2,\cdots,x_n<\theta+1/2}$，

即 $\theta-1/2<x_{(1)}\le x_{(n)}<\theta+1/2$，可得当 $x_{(n)}-1/2<\theta<x_{(1)}+1/2$ 时，都有 $L(\theta)=1$，

故 θ 的最大似然估计 $\hat{\theta}$ 是 $(x_{(n)}-1/2, x_{(1)}+1/2)$ 中任何一个值；

（3）因 $L(\theta_1,\theta_2)=\prod_{i=1}^n\dfrac{1}{\theta_2-\theta_1}\mathrm{I}_{\theta_1<x_i<\theta_2}=\dfrac{1}{(\theta_2-\theta_1)^n}\mathrm{I}_{\theta_1<x_1,x_2,\cdots,x_n<\theta_2}$，

显然 θ_1 越大且 θ_2 越小时，$L(\theta_1,\theta_2)$ 越大，但只有 $\theta_1<x_1, x_2, \cdots, x_n<\theta_2$ 时，才有 $L(\theta_1,\theta_2)>0$，

即 $\theta_1=\min\{x_1, x_2, \ldots, x_n\}$ 且 $\theta_2=\max\{x_1, x_2, \ldots, x_n\}$ 时，$L(\theta_1,\theta_2)$ 达到最大，

故 θ_1 的最大似然估计 $\hat{\theta}_1=X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$，

θ_2 的最大似然估计 $\hat{\theta}_2=X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}$。

4． 一地质学家为研究密歇根湖的湖滩地区的岩石成分，随机地自该地区取 100 个样品，每个样品有 10 块石子，记录了每个样品中属石灰石的石子数。假设这 100 次观察相互独立，求这地区石子中石灰石的比例 p 的最大似然估计。该地质学家所得的数据如下：

样本中的石子数	0	1	2	3	4	5	6	7	8	9	10
样品个数	0	1	6	7	23	26	21	12	3	1	0

解：总体 X 为样品的 10 块石子中属石灰石的石子数，即 X 服从二项分布 $B(10, p)$，其概率函数为

$$p(x)=\binom{10}{x}p^x(1-p)^{10-x}\text{，}x=1, 2, \cdots, 10\text{，}$$

因 $L(p)=\prod_{i=1}^{n}\binom{10}{x_i}p^{x_i}(1-p)^{10-x_i}=\prod_{i=1}^{100}\binom{10}{x_i}\cdot p^{\sum_{i=1}^{100}x_i}(1-p)^{1000-\sum_{i=1}^{100}x_i}$，

即 $\ln L(p)=\sum_{i=1}^{100}\ln\binom{10}{x_i}+\sum_{i=1}^{100}x_i\cdot\ln p+\left(1000-\sum_{i=1}^{100}x_i\right)\cdot\ln(1-p)$，

令 $\dfrac{d\ln L(p)}{dp}=\sum_{i=1}^{100}x_i\cdot\dfrac{1}{p}-\left(1000-\sum_{i=1}^{100}x_i\right)\cdot\dfrac{1}{1-p}=0$，得 $p=\dfrac{1}{1000}\sum_{i=1}^{100}x_i$，即 $\hat{p}=\dfrac{1}{1000}\sum_{i=1}^{100}X_i$

由于 $\sum_{i=1}^{100}x_i=0+1\times1+6\times2+7\times3+1\times9+0=499$，

故比例 p 的最大似然估计 $\hat{p}=\dfrac{1}{1000}\times499=0.499$。

5． 在遗传学研究中经常要从截尾二项分布中抽样，其总体概率函数为

$$P\{X=k; p\}=\frac{\binom{m}{k}p^k(1-p)^{m-k}}{1-(1-p)^m}\text{，}\quad k=1, 2, \cdots, m\text{。}$$

若已知 $m=2$，$X_1, \cdots, X_n$ 是样本，试求 p 的最大似然估计。

解：当 $m=2$ 时，X 只能取值 1 或 2，且 $P\{X=1\}=\dfrac{2p(1-p)}{1-(1-p)^2}=\dfrac{2-2p}{2-p}$，$P\{X=2\}=\dfrac{p^2}{1-(1-p)^2}=\dfrac{p}{2-p}$，

即 $P\{X=x; p\}=\left(\dfrac{2-2p}{2-p}\right)^{2-x}\left(\dfrac{p}{2-p}\right)^{x-1}=\dfrac{(2-2p)^{2-x}p^{x-1}}{2-p}$，$x=1, 2$，

因 $L(p)=\prod_{i=1}^{n}\dfrac{(2-2p)^{2-x_i}p^{x_i-1}}{2-p}=\dfrac{(2-2p)^{2n-\sum_{i=1}^{n}x_i}p^{\sum_{i=1}^{n}x_i-n}}{(2-p)^n}$，

即 $\ln L(p)=\left(2n-\sum_{i=1}^{n}x_i\right)\cdot\ln(2-2p)+\left(\sum_{i=1}^{n}x_i-n\right)\cdot\ln p-n\ln(2-p)$，

令 $\dfrac{d\ln L(p)}{dp}=\left(2n-\sum_{i=1}^{n}x_i\right)\cdot\dfrac{-2}{2-2p}+\left(\sum_{i=1}^{n}x_i-n\right)\cdot\dfrac{1}{p}-n\cdot\dfrac{-1}{2-p}=0$，得 $p=2-\dfrac{2n}{\sum_{i=1}^{n}x_i}=2-\dfrac{2}{\bar{x}}$，

故 p 的最大似然估计 $\hat{p}=2-\dfrac{2}{\overline{X}}$ 。

6． 已知在文学家萧伯纳的“An Intelligent Woman’s Guide to Socialism”一书中，一个句子的单词数 X 近似地服从对数正态分布，即 $Z=\ln X\sim N(\mu,\sigma^2)$。今从该书中随机地取 20 个句子，这些句子中的单词数分别为

52, 24, 15, 67, 15, 22, 63, 26, 16, 32, 7, 33, 28, 14, 7, 29, 10, 6, 59, 30，

求该书中一个句子单词数均值 $E(X)=\mathrm{e}^{\mu+\sigma^2/2}$ 的最大似然估计。

解：因 $Z=\ln X\sim N(\mu,\sigma^2)$，

则 μ 的最大似然估计 $\hat{\mu}=\bar{z}=\dfrac{1}{n}\sum\limits_{i=1}^{n}z_i=\dfrac{1}{n}\sum\limits_{i=1}^{n}\ln x_i=\dfrac{1}{20}(\ln 52+\ln 24+\cdots+\ln 30)=3.09$,

σ^2 的最大似然估计

$$\hat{\sigma^2}=s_z^{*2}=\frac{1}{n}\sum_{i=1}^{n}(z_i-\bar{z})^2=\frac{1}{20}[(\ln 52-3.09)^2+(\ln 24-3.09)^2+\cdots+(\ln 30-3.09)^2]=0.51,$$

故由最大似然估计的不变性知 $E(X)=\mathrm{e}^{\mu+\sigma^2/2}$ 的最大似然估计 $\widehat{E(X)}=\mathrm{e}^{\bar{z}+s_z^{*2}/2}=\mathrm{e}^{3.09+0.51/2}=28.31$ 。

7． 总体 $X\sim U(\theta,2\theta)$，其中 $\theta>0$ 是未知参数，又 $X_1,\cdots,X_n$ 为取自该总体的样本，$\overline{X}$ 为样本均值。

（1）证明 $\hat{\theta}=\dfrac{2}{3}\overline{X}$ 是参数 θ 的无偏估计和相合估计；

（2）求 θ 的最大似然估计，它是无偏估计吗？是相合估计吗？

解：（1）因 $X\sim U(\theta,2\theta)$，有 $E(X)=\dfrac{\theta+2\theta}{2}=\dfrac{3}{2}\theta$， $\mathrm{Var}(X)=\dfrac{(2\theta-\theta)^2}{12}=\dfrac{1}{12}\theta^2$,

故 $E(\hat{\theta})=\dfrac{2}{3}E(\overline{X})=\dfrac{2}{3}E(X)=\dfrac{2}{3}\cdot\dfrac{3}{2}\theta=\theta$，即 $\hat{\theta}=\dfrac{2}{3}\overline{X}$ 是参数 θ 的无偏估计；

因 $\mathrm{Var}(\hat{\theta})=\dfrac{4}{9}\mathrm{Var}(\overline{X})=\dfrac{4}{9n}\mathrm{Var}(X)=\dfrac{4}{9n}\cdot\dfrac{1}{12}\theta^2=\dfrac{\theta^2}{27n}$，有 $\lim\limits_{n\to\infty}E(\hat{\theta})=\theta$， $\lim\limits_{n\to\infty}\mathrm{Var}(\hat{\theta})=0$,

故 $\hat{\theta}=\dfrac{2}{3}\overline{X}$ 是参数 θ 的相合估计；

（2）因 $L(\theta)=\prod\limits_{i=1}^{n}\dfrac{1}{\theta}\mathrm{I}_{\theta<x_i<2\theta}=\dfrac{1}{\theta^n}\mathrm{I}_{\theta<x_1,x_2,\cdots,x_n<2\theta}$,

显然 θ 越小，$\dfrac{1}{\theta^n}$ 越大，但只有 $\theta<x_1,x_2,\cdots,x_n<2\theta$ 时，才有 $L(\theta)>0$，

即 $\theta=\dfrac{1}{2}\max\{x_1,x_2,\cdots,x_n\}$ 时，$L(\theta)$ 达到最大，

故 θ 的最大似然估计为 $\hat{\theta}^*=\dfrac{1}{2}X_{(n)}=\dfrac{1}{2}\max\{X_1,X_2,\cdots,X_n\}$；

因 X 的密度函数为 $p(x)=\begin{cases}\dfrac{1}{\theta}, & \theta<x<2\theta;\\ 0, & \text{其他}.\end{cases}$，分布函数为 $F(x)=\begin{cases}0, & x<\theta;\\ \dfrac{x-\theta}{\theta}, & \theta\le x<2\theta;\\ 1, & x\ge 2\theta.\end{cases}$

则 $X_{(n)}$ 的密度函数 $p_n(x)=n[F(x)]^{n-1}p(x)=\begin{cases}\dfrac{n(x-\theta)^{n-1}}{\theta^n}, & \theta<x<2\theta;\\ 0, & \text{其他}.\end{cases}$

因 $E(X_{(n)}-\theta)=\int_\theta^{2\theta}(x-\theta)\cdot\frac{n(x-\theta)^{n-1}}{\theta^n}dx=\frac{n}{\theta^n}\cdot\frac{(x-\theta)^{n+1}}{n+1}\bigg|_\theta^{2\theta}=\frac{n}{n+1}\theta$，有 $E(X_{(n)})=\frac{2n+1}{n+1}\theta$，

且 $E[(X_{(n)}-\theta)^2]=\int_\theta^{2\theta}(x-\theta)^2\cdot\frac{n(x-\theta)^{n-1}}{\theta^n}dx=\frac{n}{\theta^n}\cdot\frac{(x-\theta)^{n+2}}{n+2}\bigg|_\theta^{2\theta}=\frac{n}{n+2}\theta^2$，

则 $\mathrm{Var}(X_{(n)})=\mathrm{Var}(X_{(n)}-\theta)=\frac{n}{n+2}\theta^2-\left(\frac{n}{n+1}\theta\right)^2=\frac{n}{(n+1)^2(n+2)}\theta^2$，

因 $E(\hat{\theta}^*)=\frac{1}{2}E(X_{(n)})=\frac{2n+1}{2(n+1)}\theta\neq\theta$，$\mathrm{Var}(\hat{\theta}^*)=\frac{1}{4}\mathrm{Var}(X_{(n)})=\frac{n}{4(n+1)^2(n+2)}\theta^2$，

故 $\hat{\theta}^*=\frac{1}{2}X_{(n)}$ 不是参数 θ 的无偏估计，应该修偏为 $\hat{\theta}=\frac{n+1}{2n+1}X_{(n)}$ 才是 θ 的无偏估计，

因 $\lim\limits_{n\to\infty}E(\hat{\theta}^*)=\lim\limits_{n\to\infty}\frac{2n+1}{2(n+1)}\theta=\theta$，$\lim\limits_{n\to\infty}\mathrm{Var}(\hat{\theta}^*)=\lim\limits_{n\to\infty}\frac{n}{4(n+1)^2(n+2)}\theta^2=0$，

故 θ 的最大似然估计 $\hat{\theta}^*=\frac{1}{2}X_{(n)}$ 是参数 θ 的相合估计。

8．设 $X_1,\cdots,X_n$ 是来自密度函数为 $p(x;\theta)=\mathrm{e}^{-(x-\theta)}, x>\theta$ 的样本。

（1）求 θ 的最大似然估计 $\hat{\theta}_1$，它是否是相合估计？是否是无偏估计？

（2）求 θ 的矩估计 $\hat{\theta}_2$，它是否是相合估计？是否是无偏估计？

解：（1）似然函数 $L(\theta)=\prod_{i=1}^n\mathrm{e}^{-(x_i-\theta)}\mathrm{I}_{x_i>\theta}=\mathrm{e}^{-\sum_{i=1}^n x_i+n\theta}\mathrm{I}_{x_1,x_2,\cdots,x_n>\theta}$，

显然 θ 越大，$\mathrm{e}^{-\sum_{i=1}^n x_i+n\theta}$ 越大，但只有 $x_1,x_2,\cdots,x_n>\theta$ 时，才有 $L(\theta)>0$，
即 $\theta=\min\{x_1,x_2,\ldots,x_n\}$ 时，$L(\theta)$ 达到最大，

故 θ 的最大似然估计 $\hat{\theta}_1=X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$；

因 X 的密度函数与分布函数分别为

$$p(x)=\begin{cases}\mathrm{e}^{-(x-\theta)}, & x>\theta;\\ 0, & x\le\theta.\end{cases}\quad F(x)=\begin{cases}1-\mathrm{e}^{-(x-\theta)}, & x>\theta;\\ 0, & x\le\theta.\end{cases}$$

则 $X_{(1)}$ 的密度函数为

$$p_1(x)=n[1-F(x)]^{n-1}p(x)=\begin{cases}n\mathrm{e}^{-n(x-\theta)}, & x>\theta;\\ 0, & x\le\theta.\end{cases}$$

可得 $X_{(1)}-\theta$ 服从指数分布 $Exp(n)$，

因 $E(X_{(1)}-\theta)=\frac{1}{n}$，$\mathrm{Var}(X_{(1)}-\theta)=\frac{1}{n^2}$，

则 $E(\hat{\theta}_1)=E(X_{(1)})=\theta+\frac{1}{n}\neq\theta$，$\mathrm{Var}(\hat{\theta}_1)=\mathrm{Var}(X_{(1)})=\mathrm{Var}(X_{(1)}-\theta)=\frac{1}{n^2}$，

故 $\hat{\theta}_1=X_{(1)}$ 不是 θ 的无偏估计；

因 $\lim\limits_{n\to\infty} E(\hat{\theta}_1)=\lim\limits_{n\to\infty}\left(\theta+\dfrac{1}{n}\right)=\theta$， $\lim\limits_{n\to\infty}\text{Var}(\hat{\theta}_1)=\lim\limits_{n\to\infty}\dfrac{1}{n^2}=0$，

故 $\hat{\theta}_1=X_{(1)}$ 是 θ 的相合估计；

（2）因总体 X 的密度函数为 $p(x;\theta)=\mathrm{e}^{-(x-\theta)}, x>\theta$， 有 $X-\theta$ 服从指数分布 $Exp(1)$，
则 $E(X-\theta)=E(X)-\theta=1$，即 $\theta=E(X)-1$，

故 θ 的矩估计 $\hat{\theta}_2=\bar{X}-1$；

因 $E(X)=\theta+1$，$\text{Var}(X)=\text{Var}(X-\theta)=1$，

则 $E(\hat{\theta}_2)=E(\bar{X})-1=E(X)-1=\theta$， $\text{Var}(\hat{\theta}_2)=\text{Var}(\bar{X})=\dfrac{1}{n}\text{Var}(X)=\dfrac{1}{n}$，

故 $\hat{\theta}_2=\bar{X}-1$ 是 θ 的无偏估计；

因 $\lim\limits_{n\to\infty} E(\hat{\theta}_2)=\theta$， $\lim\limits_{n\to\infty}\text{Var}(\hat{\theta}_2)=\lim\limits_{n\to\infty}\dfrac{1}{n}=0$，

故 $\hat{\theta}_2=\bar{X}-1$ 是 θ 的相合估计。

9． 设总体 $X\sim Exp(1/\theta)$，$X_1, \cdots, X_n$ 是样本，θ 的矩估计和最大似然估计都是 $\bar{X}$，它也是 θ 的相合估计和无偏估计，试证明在均方误差准则下存在优于 $\bar{X}$ 的估计（提示：考虑 $\hat{\theta}_a=a\bar{X}$，找均方误差最小者）。

证：因 $X\sim Exp(1/\theta)$，有 $E(X)=\theta$，$\text{Var}(X)=\theta^2$，且 X 的密度函数为

$$p(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, & x>0;\\ 0, & x\le 0.\end{cases}$$

故 $\theta=E(X)$，即 θ 的矩估计为 $\hat{\theta}=\bar{X}$；

因似然函数 $L(\theta)=\prod\limits_{i=1}^{n}\dfrac{1}{\theta}\mathrm{e}^{-\frac{x_i}{\theta}}\mathrm{I}_{x_i>0}=\dfrac{1}{\theta^n}\mathrm{e}^{-\frac{1}{\theta}\sum\limits_{i=1}^{n}x_i}\mathrm{I}_{x_1,x_2,\cdots,x_n>0}$，

当 $x_1, x_2, \cdots, x_n>0$ 时， $\ln L(\theta)=-n\ln\theta-\dfrac{1}{\theta}\sum\limits_{i=1}^{n}x_i$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=-\dfrac{n}{\theta}+\dfrac{1}{\theta^2}\sum\limits_{i=1}^{n}x_i=0$，得 $\theta=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}$，

故 θ 的最大似然估计也为 $\hat{\theta}=\bar{X}$；

因 $E(\bar{X})=E(X)=\theta$， $\text{Var}(\bar{X})=\dfrac{1}{n}\text{Var}(X)=\dfrac{\theta^2}{n}$，

故 $\bar{X}$ 是 θ 的无偏估计；

因 $\lim\limits_{n\to\infty} E(\bar{X})=\theta$， $\lim\limits_{n\to\infty}\text{Var}(\bar{X})=\lim\limits_{n\to\infty}\dfrac{\theta^2}{n}=0$，

故 $\bar{X}$ 是 θ 的相合估计；

设 $\hat{\theta}_a=a\bar{X}$，有 $E(\hat{\theta}_a)=aE(\bar{X})=a\theta$， $\text{Var}(\hat{\theta}_a)=a^2\,\text{Var}(\bar{X})=\dfrac{a^2\theta^2}{n}$，

则 $\text{MSE}(\bar{X})=\text{Var}(\bar{X})+[E(\bar{X})-\theta]^2=\dfrac{\theta^2}{n}+(\theta-\theta)^2=\dfrac{\theta^2}{n}$，

$$\text{MSE}(\hat{\theta}_a)=\text{Var}(\hat{\theta}_a)+[E(\hat{\theta}_a)-\theta]^2=\frac{a^2\theta^2}{n}+(a\theta-\theta)^2=\left(\frac{a^2}{n}+a^2-2a+1\right)\theta^2$$

$$=\left(\frac{n+1}{n}a^2-2a+\frac{n}{n+1}+\frac{1}{n+1}\right)\theta^2=\left[\frac{n+1}{n}\left(a-\frac{n}{n+1}\right)^2+\frac{1}{n+1}\right]\theta^2,$$

故当 $a=\dfrac{n}{n+1}$ 时，$\hat{\theta}_a=\dfrac{n}{n+1}\bar{X}$ 的均方误差 $\text{MSE}(\hat{\theta}_a)=\dfrac{\theta^2}{n+1}$ 小于 $\bar{X}$ 的均方误差 $\text{MSE}(\bar{X})=\dfrac{\theta^2}{n}$。

10．为了估计湖中有多少条鱼，从中捞出 1000 条，标上记号后放回湖中，然后再捞出 150 条鱼，发现其中有 10 条鱼有记号。问湖中有多少条鱼，才能使 150 条鱼中出现 10 条带记号的鱼的概率最大？

解：设湖中有 N 条鱼，有湖中每条鱼带记号的概率为 $p=\dfrac{1000}{N}$，

看作总体 X 服从两点分布 $b(1, p)$，从中抽取容量为 150 的样本 $X_1, X_2, \cdots, X_{150}$，有 $\sum\limits_{i=1}^{150}x_i=10$，

似然函数 $L(p)=\prod\limits_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}$，有 $\ln L(p)=\sum\limits_{i=1}^{n}x_i\cdot\ln p+\left(n-\sum\limits_{i=1}^{n}x_i\right)\cdot\ln(1-p)$，

令 $\dfrac{d\ln L(p)}{dp}=\sum\limits_{i=1}^{n}x_i\cdot\dfrac{1}{p}+\left(n-\sum\limits_{i=1}^{n}x_i\right)\cdot\dfrac{-1}{1-p}=0$，得 $p=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}$，即 p 的最大似然估计为 $\hat{p}=\bar{X}$，

因 $N=\dfrac{1000}{p}$，由最大似然估计的不变性知 $\hat{N}=\dfrac{1000}{\bar{X}}$，

故湖中有 $\hat{N}=\dfrac{1000}{\frac{1}{150}\times 10}=15000$ 条鱼时，才能使 150 条鱼中出现 10 条带记号的鱼的概率最大。

11．证明：对正态分布 $N(\mu, \sigma^2)$，若只有一个观测值，则 μ, σ^2 的最大似然估计不存在。

证：若只有一个观测值，似然函数 $L(\mu,\sigma^2)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$，

对于任一固定的 σ，当 $\mu=x$ 时，$L(\mu)$ 取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$，

但显然 σ 越小，$\dfrac{1}{\sqrt{2\pi}\sigma}$ 越大，且 σ 可任意接近于 0，即 $\dfrac{1}{\sqrt{2\pi}\sigma}$ 不存在最大值，

故 μ, σ^2 的最大似然估计不存在。