# 第六章　参数估计

## 习题 6.1

1． 设 $X_1, X_2, X_3$ 是取自某总体容量为 3 的样本，试证下列统计量都是该总体均值$\mu$ 的无偏估计，在方差存在时指出哪一个估计的有效性最差？

（1）$\hat{\mu}_1=\dfrac{1}{2}X_1+\dfrac{1}{3}X_2+\dfrac{1}{6}X_3$；　（2）$\hat{\mu}_2=\dfrac{1}{3}X_1+\dfrac{1}{3}X_2+\dfrac{1}{3}X_3$；　（3）$\hat{\mu}_3=\dfrac{1}{6}X_1+\dfrac{1}{6}X_2+\dfrac{2}{3}X_3$．

证：因 $E(\hat{\mu}_1)=\dfrac{1}{2}E(X_1)+\dfrac{1}{3}E(X_2)+\dfrac{1}{6}E(X_3)=\dfrac{1}{2}\mu+\dfrac{1}{3}\mu+\dfrac{1}{6}\mu=\mu$，

$$E(\hat{\mu}_2)=\frac{1}{3}E(X_1)+\frac{1}{3}E(X_2)+\frac{1}{3}E(X_3)=\frac{1}{3}\mu+\frac{1}{3}\mu+\frac{1}{3}\mu=\mu,$$

$$E(\hat{\mu}_3)=\frac{1}{6}E(X_1)+\frac{1}{6}E(X_2)+\frac{2}{3}E(X_3)=\frac{1}{6}\mu+\frac{1}{6}\mu+\frac{2}{3}\mu=\mu,$$

故 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是总体均值$\mu$ 的无偏估计；

因 $\mathrm{Var}(\hat{\mu}_1)=\dfrac{1}{4}\mathrm{Var}(X_1)+\dfrac{1}{9}\mathrm{Var}(X_2)+\dfrac{1}{36}\mathrm{Var}(X_3)=\dfrac{1}{4}\sigma^2+\dfrac{1}{9}\sigma^2+\dfrac{1}{36}\sigma^2=\dfrac{14}{36}\sigma^2$，

$$\mathrm{Var}(\hat{\mu}_2)=\frac{1}{9}\mathrm{Var}(X_1)+\frac{1}{9}\mathrm{Var}(X_2)+\frac{1}{9}\mathrm{Var}(X_3)=\frac{1}{9}\sigma^2+\frac{1}{9}\sigma^2+\frac{1}{9}\sigma^2=\frac{1}{3}\sigma^2,$$

$$\mathrm{Var}(\hat{\mu}_3)=\frac{1}{36}\mathrm{Var}(X_1)+\frac{1}{36}\mathrm{Var}(X_2)+\frac{4}{9}\mathrm{Var}(X_3)=\frac{1}{36}\sigma^2+\frac{1}{36}\sigma^2+\frac{4}{9}\sigma^2=\frac{1}{2}\sigma^2,$$

故 $\mathrm{Var}(\hat{\mu}_2)<\mathrm{Var}(\hat{\mu}_1)<\mathrm{Var}(\hat{\mu}_3)$，即 $\hat{\mu}_2$ 有效性最好，$\hat{\mu}_1$ 其次，$\hat{\mu}_3$ 最差．

2． 设 $X_1, X_2, \cdots, X_n$ 是来自 $Exp(\lambda)$的样本，已知 $\overline{X}$ 为 $1/\lambda$的无偏估计，试说明$1/\overline{X}$ 是否为$\lambda$的无偏估计．

解：因 $X_1, X_2, \cdots, X_n$ 相互独立且都服从指数分布 $Exp(\lambda)$，即都服从伽玛分布 $Ga(1, \lambda)$，

由伽玛分布的可加性知 $Y=\sum\limits_{i=1}^{n}X_i$ 服从伽玛分布 $Ga(n, \lambda)$，密度函数为

$$p_Y(y)=\frac{\lambda^n}{\Gamma(n)}y^{n-1}\mathrm{e}^{-\lambda y}\mathrm{I}_{y>0},$$

则 $E\left(\dfrac{1}{\overline{X}}\right)=E\left(\dfrac{n}{Y}\right)=\displaystyle\int_0^{+\infty}\frac{n}{y}\cdot\frac{\lambda^n}{\Gamma(n)}y^{n-1}\mathrm{e}^{-\lambda y}dy=\frac{n\lambda^n}{\Gamma(n)}\int_0^{+\infty}y^{n-2}\mathrm{e}^{-\lambda y}dy=\frac{n\lambda^n}{\Gamma(n)}\cdot\frac{\Gamma(n-1)}{\lambda^{n-1}}=\frac{n}{n-1}\lambda$，

故$1/\overline{X}$ 不是$\lambda$的无偏估计．

3． 设 $\hat{\theta}$ 是参数$\theta$ 的无偏估计，且有 $\mathrm{Var}(\hat{\theta})>0$，试证 $(\hat{\theta})^2$ 不是$\theta^2$ 的无偏估计．

证：因 $E(\hat{\theta})=\theta$，有 $E[(\hat{\theta})^2]=\mathrm{Var}(\hat{\theta})+[E(\hat{\theta})]^2=\mathrm{Var}(\hat{\theta})+\theta^2>\theta^2$，故 $(\hat{\theta})^2$ 不是$\theta^2$ 的无偏估计．

4． 设总体 $X\sim N(\mu, \sigma^2)$，$X_1, \cdots, X_n$ 是来自该总体的一个样本．试确定常数 $c$ 使 $c\sum\limits_{i=1}^{n}(X_{i+1}-X_i)^2$ 为$\sigma^2$ 的无偏估计．

解：因 $E[(X_{i+1}-X_i)^2]=\mathrm{Var}(X_{i+1}-X_i)+[E(X_{i+1}-X_i)]^2=\mathrm{Var}(X_{i+1})+\mathrm{Var}(X_i)+[E(X_{i+1})-E(X_i)]^2=2\sigma^2$，

则$E\left[c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right]=c\sum_{i=1}^{n-1}E[(X_{i+1}-X_i)^2]=c\cdot(n-1)\cdot 2\sigma^2=2c(n-1)\sigma^2$，

故当$c=\dfrac{1}{2(n-1)}$时，$E\left[c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2\right]=\sigma^2$，即$c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$是$\sigma^2$的无偏估计.

5． 设$X_1, X_2, \cdots, X_n$是来自下列总体中抽取的简单样本，

$$p(x;\theta)=\begin{cases}1, & \theta-\dfrac{1}{2}\le x\le\theta+\dfrac{1}{2};\\ 0, & \text{其他}.\end{cases}$$

证明样本均值$\overline{X}$及$\dfrac{1}{2}(X_{(1)}+X_{(n)})$都是$\theta$的无偏估计，问何者更有效？

证：因总体$X\sim U\left(\theta-\dfrac{1}{2},\theta+\dfrac{1}{2}\right)$，有$Y=X-\theta+\dfrac{1}{2}\sim U(0,1)$，

则$\overline{X}=\overline{Y}+\theta-\dfrac{1}{2}$，$X_{(1)}=Y_{(1)}+\theta-\dfrac{1}{2}$，$X_{(n)}=Y_{(n)}+\theta-\dfrac{1}{2}$，即$\dfrac{1}{2}(X_{(1)}+X_{(n)})=\dfrac{1}{2}(Y_{(1)}+Y_{(n)})+\theta-\dfrac{1}{2}$，

可得$E(\overline{X})=E(\overline{Y})+\theta-\dfrac{1}{2}=E(Y)+\theta-\dfrac{1}{2}=\theta$，$\mathrm{Var}(\overline{X})=\mathrm{Var}(\overline{Y})=\dfrac{1}{n}\mathrm{Var}(Y)=\dfrac{1}{12n}$，

因$Y$的密度函数与分布函数分别为

$$p_Y(y)=\mathrm{I}_{0<y<1},\quad F_Y(y)=\begin{cases}0, & y<0;\\ y, & 0\le y<1;\\ 1, & y\ge 1.\end{cases}$$

有$Y_{(1)}$与$Y_{(n)}$的密度函数分别为

$$p_1(y)=n[1-F_Y(y)]^{n-1}p_Y(y)=n(1-y)^{n-1}\mathrm{I}_{0<y<1},\quad p_n(y)=n[F_Y(y)]^{n-1}p_Y(y)=ny^{n-1}\mathrm{I}_{0<y<1},$$

且$(Y_{(1)}, Y_{(n)})$的联合密度函数为

$$p_{1n}(y_{(1)},y_{(n)})=n(n-1)[F_Y(y_{(n)})-F_Y(y_{(1)})]^{n-2}p_Y(y_{(1)})p_Y(y_{(n)})\mathrm{I}_{y_{(1)}<y_{(n)}}$$

$$=n(n-1)(y_{(n)}-y_{(1)})^{n-2}\mathrm{I}_{0<y_{(1)}<y_{(n)}<1},$$

则$E(Y_{(1)})=\int_0^1 y\cdot n(1-y)^{n-1}dy=n\cdot\dfrac{\Gamma(2)\Gamma(n)}{\Gamma(2+n)}=\dfrac{1}{n+1}$，$E(Y_{(n)})=\int_0^1 y\cdot ny^{n-1}dy=\dfrac{n}{n+1}$，

$E(Y_{(1)}^2)=\int_0^1 y^2\cdot n(1-y)^{n-1}dy=n\cdot\dfrac{\Gamma(3)\Gamma(n)}{\Gamma(3+n)}=\dfrac{2}{(n+1)(n+2)}$，$E(Y_{(n)}^2)=\int_0^1 y^2\cdot ny^{n-1}dy=\dfrac{n}{n+2}$，

$$E(Y_{(1)}Y_{(n)})=\int_0^1 dy_{(n)}\int_0^{y_{(n)}}y_{(1)}y_{(n)}\cdot n(n-1)(y_{(n)}-y_{(1)})^{n-2}dy_{(1)}=\int_0^1 dy_{(n)}\int_0^{y_{(n)}}y_{(1)}y_{(n)}\cdot n\cdot(-1)d(y_{(n)}-y_{(1)})^{n-1}$$

$$=\int_0^1 dy_{(n)}\left[-ny_{(1)}y_{(n)}(y_{(n)}-y_{(1)})^{n-1}\Big|_0^{y_{(n)}}+\int_0^{y_{(n)}}n(y_{(n)}-y_{(1)})^{n-1}\cdot y_{(n)}dy_{(1)}\right]$$

$$=\int_0^1 dy_{(n)}\left[-y_{(n)}\cdot(y_{(n)}-y_{(1)})^n\Big|_0^{y_{(n)}}\right]=\int_0^1 y_{(n)}^{n+1}dy_{(n)}=\frac{1}{n+2}y_{(n)}^{n+2}\Big|_0^1=\frac{1}{n+2},$$

即$\mathrm{Var}(Y_{(1)})=\dfrac{2}{(n+1)(n+2)}-\left(\dfrac{1}{n+1}\right)^2=\dfrac{n}{(n+1)^2(n+2)}$，$\mathrm{Var}(Y_{(n)})=\dfrac{n}{n+2}-\left(\dfrac{n}{n+1}\right)^2=\dfrac{n}{(n+1)^2(n+2)}$，

且$\mathrm{Cov}(Y_{(1)},Y_{(n)})=\dfrac{1}{n+2}-\dfrac{1}{n+1}\cdot\dfrac{n}{n+1}=\dfrac{1}{(n+1)^2(n+2)}$

可得$E\left[\dfrac{1}{2}(X_{(1)}+X_{(n)})\right]=\dfrac{1}{2}[E(Y_{(1)})+E(Y_{(n)})]+\theta-\dfrac{1}{2}=\theta$，

$$\mathrm{Var}\left[\frac{1}{2}(X_{(1)}+X_{(n)})\right]=\frac{1}{4}[\mathrm{Var}(Y_{(1)})+\mathrm{Var}(Y_{(n)})+2\,\mathrm{Cov}(Y_{(1)},Y_{(n)})]=\frac{2n+2}{4(n+1)^2(n+2)}=\frac{1}{2(n+1)(n+2)},$$

因$E(\bar{X})=\theta$，$E\left[\dfrac{1}{2}(X_{(1)}+X_{(n)})\right]=\theta$，

故$\bar{X}$及$\dfrac{1}{2}(X_{(1)}+X_{(n)})$都是$\theta$的无偏估计；

因当$n>1$时，$\mathrm{Var}(\bar{X})=\dfrac{1}{12n}>\mathrm{Var}\left[\dfrac{1}{2}(X_{(1)}+X_{(n)})\right]=\dfrac{1}{2(n+1)(n+2)}$，

故$\dfrac{1}{2}(X_{(1)}+X_{(n)})$比样本均值$\bar{X}$更有效.

6．设$X_1, X_2, X_3$服从均匀分布$U(0,\theta)$，试证$\dfrac{4}{3}X_{(3)}$及$4X_{(1)}$都是$\theta$的无偏估计量，哪个更有效？

解：因总体$X$的密度函数与分布函数分别为

$$p(x)=\frac{1}{\theta}\mathrm{I}_{0<x<\theta}，\quad F(x)=\begin{cases}0, & x<0;\\ \dfrac{x}{\theta}, & 0\le x<\theta;\\ 1, & x\ge\theta.\end{cases}$$

有$X_{(1)}$与$X_{(3)}$的密度函数分别为

$$p_1(x)=3[1-F(x)]^2p(x)=\frac{3(\theta-x)^2}{\theta^3}\mathrm{I}_{0<x<\theta}，\quad p_3(x)=3[F(x)]^2p(x)=\frac{3x^2}{\theta^3}\mathrm{I}_{0<x<\theta}，$$

则$E(X_{(1)})=\displaystyle\int_0^\theta x\cdot\frac{3(\theta-x)^2}{\theta^3}dx=\frac{3}{\theta^3}\left(\theta^2\cdot\frac{x^2}{2}-2\theta\cdot\frac{x^3}{3}+\frac{x^4}{4}\right)\Bigg|_0^\theta=\frac{\theta}{4}$，

$$E(X_{(3)})=\int_0^\theta x\cdot\frac{3x^2}{\theta^3}dy=\frac{3}{\theta^3}\cdot\frac{x^4}{4}\bigg|_0^\theta=\frac{3\theta}{4},$$

$$E(X_{(1)}^2)=\int_0^\theta x^2\cdot\frac{3(\theta-x)^2}{\theta^3}dx=\frac{3}{\theta^3}\left(\theta^2\cdot\frac{x^3}{3}-2\theta\cdot\frac{x^4}{4}+\frac{x^5}{5}\right)\Bigg|_0^\theta=\frac{\theta^2}{10},$$

$$E(X_{(3)}^2)=\int_0^\theta x^2\cdot\frac{3x^2}{\theta^3}dy=\frac{3}{\theta^3}\cdot\frac{x^5}{5}\bigg|_0^\theta=\frac{3\theta^2}{5},$$

即$\mathrm{Var}(X_{(1)})=\dfrac{\theta^2}{10}-\left(\dfrac{\theta}{4}\right)^2=\dfrac{3\theta^2}{80}$，$\mathrm{Var}(X_{(3)})=\dfrac{3\theta^2}{5}-\left(\dfrac{3\theta}{4}\right)^2=\dfrac{3\theta^2}{80}$，

因$E(4X_{(1)})=4\cdot\dfrac{\theta}{4}=\theta$，$E\left(\dfrac{4}{3}X_{(3)}\right)=\dfrac{4}{3}\cdot\dfrac{3\theta}{4}=\theta$，

故 $4X_{(1)}$及 $\frac{4}{3}X_{(3)}$ 都是 $\theta$ 的无偏估计；

因 $\text{Var}(4X_{(1)})=16\cdot\frac{3\theta^2}{80}=\frac{3\theta^2}{5}$， $\text{Var}\left(\frac{4}{3}X_{(3)}\right)=\frac{16}{9}\cdot\frac{3\theta^2}{80}=\frac{\theta^2}{15}$，有 $\text{Var}(4X_{(1)})>\text{Var}\left(\frac{4}{3}X_{(3)}\right)$，

故 $\frac{4}{3}X_{(3)}$ 比 $4X_{(1)}$更有效.

7. 设从均值为 $\mu$，方差为 $\sigma^2>0$ 的总体中，分别抽取容量为 $n_1$ 和 $n_2$ 的两独立样本，$\overline{X}_1$ 和 $\overline{X}_2$ 分别是这两个样本的均值．试证，对于任意常数 $a, b$（$a+b=1$），$Y=a\overline{X}_1+b\overline{X}_2$ 都是 $\mu$ 的无偏估计，并确定常数 $a, b$ 使 $\text{Var}(Y)$ 达到最小.

解：因 $E(Y)=aE(\overline{X}_1)+bE(\overline{X}_2)=a\mu+b\mu=(a+b)\mu=\mu$，

故 $Y$ 是 $\mu$ 的无偏估计；

因 $\text{Var}(Y)=a^2\,\text{Var}(\overline{X}_1)+b^2\,\text{Var}(\overline{X}_2)=a^2\cdot\frac{\sigma^2}{n_1}+(1-a)^2\cdot\frac{\sigma^2}{n_2}=\left(\frac{n_1+n_2}{n_1n_2}a^2-\frac{2}{n_2}a+\frac{1}{n_2}\right)\sigma^2$，

令 $\frac{d}{da}\text{Var}(Y)=\left(\frac{n_1+n_2}{n_1n_2}\cdot 2a-\frac{2}{n_2}\right)\sigma^2=0$，得 $a=\frac{n_1}{n_1+n_2}$，且 $\frac{d^2}{d^2a}\text{Var}(Y)=\frac{n_1+n_2}{n_1n_2}\cdot 2\sigma^2>0$，

故当 $a=\frac{n_1}{n_1+n_2}$，$b=1-a=\frac{n_2}{n_1+n_2}$ 时，$\text{Var}(Y)$ 达到最小 $\frac{1}{n_1+n_2}\sigma^2$.

8. 设总体 $X$ 的均值为 $\mu$，方差为 $\sigma^2$，$X_1, \cdots, X_n$ 是来自该总体的一个样本，$T(X_1, \cdots, X_n)$为 $\mu$ 的任一线性无偏估计量．证明：$\overline{X}$ 与 $T$ 的相关系数为 $\sqrt{\text{Var}(\overline{X})/\text{Var}(T)}$ .

证：因 $T(X_1, \cdots, X_n)$为 $\mu$ 的任一线性无偏估计量，设 $T(X_1,\cdots,X_n)=\sum_{i=1}^{n}a_iX_i$，

则 $E(T)=\sum_{i=1}^{n}a_iE(X_i)=\mu\sum_{i=1}^{n}a_i=\mu$，即 $\sum_{i=1}^{n}a_i=1$，

因 $X_1, \cdots, X_n$ 相互独立，当 $i\ne j$ 时，有 $\text{Cov}(X_i, X_j)=0$，

则 $\text{Cov}(\overline{X},T)=\text{Cov}\left(\frac{1}{n}\sum_{i=1}^{n}X_i,\sum_{i=1}^{n}a_iX_i\right)=\sum_{i=1}^{n}\text{Cov}\left(\frac{1}{n}X_i,a_iX_i\right)=\sum_{i=1}^{n}\frac{a_i}{n}\text{Cov}(X_i,X_i)=\frac{\sigma^2}{n}\sum_{i=1}^{n}a_i=\frac{\sigma^2}{n}$，

因 $\text{Var}(\overline{X})=\frac{1}{n}\text{Var}(X)=\frac{\sigma^2}{n}=\text{Cov}(\overline{X},T)$，

故 $\overline{X}$ 与 $T$ 的相关系数为 $\text{Corr}(\overline{X},T)=\frac{\text{Cov}(\overline{X},T)}{\sqrt{\text{Var}(\overline{X})}\sqrt{\text{Var}(T)}}=\frac{\text{Var}(\overline{X})}{\sqrt{\text{Var}(\overline{X})}\sqrt{\text{Var}(T)}}=\sqrt{\frac{\text{Var}(\overline{X})}{\text{Var}(T)}}$ .

9. 设有 $k$ 台仪器，已知用第 $i$ 台仪器测量时，测定值总体的标准差为 $\sigma_i$（$i=1, \cdots, k$）．用这些仪器独立地对某一物理量 $\theta$ 各观察一次，分别得到 $X_1, \cdots, X_k$，设仪器都没有系统误差．问 $a_1, \cdots, a_k$ 应取何值，方能使 $\hat{\theta}=\sum_{i=1}^{k}a_iX_i$ 成为 $\theta$ 的无偏估计，且方差达到最小？

解：因 $E(\hat{\theta})=E\left(\sum_{i=1}^{k}a_ix_i\right)=\sum_{i=1}^{k}a_iE(x_i)=\sum_{i=1}^{k}a_i\theta=\left(\sum_{i=1}^{k}a_i\right)\theta$，

则当 $\sum_{i=1}^{k}a_i=1$ 时，$\hat{\theta}=\sum_{i=1}^{k}a_ix_i$ 是 $\theta$ 的无偏估计，

因 $\mathrm{Var}(\hat{\theta})=\mathrm{Var}\left(\sum_{i=1}^{k}a_ix_i\right)=\sum_{i=1}^{k}a_i^2\,\mathrm{Var}(x_i)=\sum_{i=1}^{k}a_i^2\sigma_i^2$，

讨论在 $\sum_{i=1}^{k}a_i=1$ 时，$\sum_{i=1}^{k}a_i^2\sigma_i^2$ 的条件极值，

设拉格朗日函数 $L(a_1,\cdots,a_k,\lambda)=\sum_{i=1}^{k}a_i^2\sigma_i^2+\lambda\left(\sum_{i=1}^{k}a_i-1\right)$，

令 $\begin{cases}\dfrac{\partial L}{\partial a_1}=2a_1\sigma_1^2+\lambda=0,\\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\ \dfrac{\partial L}{\partial a_k}=2a_k\sigma_k^2+\lambda=0,\\ \dfrac{\partial L}{\partial \lambda}=\sum_{i=1}^{k}a_i-1=0,\end{cases}$

得 $\lambda=-\dfrac{2}{\sigma_1^{-2}+\cdots+\sigma_k^{-2}}$，$a_i=\dfrac{\sigma_i^{-2}}{\sigma_1^{-2}+\cdots+\sigma_k^{-2}}$，$i=1,\cdots,k$，

故当 $a_i=\dfrac{\sigma_i^{-2}}{\sigma_1^{-2}+\cdots+\sigma_k^{-2}}$，$i=1,\cdots,k$ 时，$\hat{\theta}=\sum_{i=1}^{k}a_ix_i$ 是 $\theta$ 的无偏估计，且方差达到最小.

10．设 $X_1, X_2, \cdots, X_n$ 是来自 $N(\theta, 1)$ 的样本，证明 $g(\theta)=|\theta|$ 没有无偏估计（提示：利用 $g(\theta)$ 在 $\theta=0$ 处不可导）.

证：反证法：假设 $T=T(X_1, X_2, \cdots, X_n)$ 是 $g(\theta)=|\theta|$ 的任一无偏估计，

因 $\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$ 是 $\theta$ 的一个充分统计量，即在取定 $\overline{X}=x$ 条件下，样本条件分布与参数 $\theta$ 无关，

则 $S=E(T\,|\,\overline{X})$ 与参数 $\theta$ 无关，且 $S$ 是关于 $\overline{X}$ 的函数，$E(S)=E[E(T\,|\,\overline{X})]=E(T)=g(\theta)=|\theta|$，

可得 $S=S(\overline{X})$ 是 $g(\theta)=|\theta|$ 的无偏估计，

因 $X_1, X_2, \cdots, X_n$ 是来自 $N(\theta, 1)$ 的样本，由正态分布可加性知 $\overline{X}$ 服从正态分布 $N\left(\theta,\dfrac{1}{n}\right)$，

则 $E(S)=\int_{-\infty}^{+\infty}S(x)\cdot\dfrac{\sqrt{n}}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\sqrt{n}}{2}(x-\theta)^2}dx=\dfrac{\sqrt{n}}{\sqrt{2\pi}}\cdot\mathrm{e}^{-\frac{\sqrt{n}}{2}\theta^2}\int_{-\infty}^{+\infty}S(x)\cdot\mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}dx$，

因 $E(S)=|\theta|$，可知对任意的 $\theta$，反常积分 $\int_{-\infty}^{+\infty}\left|S(x)\right|\cdot\mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}dx$ 收敛，

则由参数$\theta$的任意性以及该反常积分在$-\infty$与$+\infty$两个方向的收敛性知$\int_{-\infty}^{+\infty}|S(x)|\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\cdot|\theta|\cdot|x|}dx$收敛，

因$\dfrac{\partial}{\partial\theta}\left[S(x)\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}\right]=S(x)\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}\cdot\sqrt{n}x$，且$|y|\le \mathrm{e}^{|y|}$，有$\left|\mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}\cdot\sqrt{n}x\right|\le \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\cdot(|\theta|+1)\cdot|x|}$，

则由$\int_{-\infty}^{+\infty}|S(x)|\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\cdot(|\theta|+1)\cdot|x|}dx$的收敛性知$\int_{-\infty}^{+\infty}\dfrac{\partial}{\partial\theta}\left[S(x)\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}\right]dx$一致收敛，

可得$E(S)=\dfrac{\sqrt{n}}{\sqrt{2\pi}}\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}\theta^2}\int_{-\infty}^{+\infty}S(x)\cdot \mathrm{e}^{-\frac{\sqrt{n}}{2}x^2+\sqrt{n}\theta x}dx$关于参数$\theta$可导，与$E(S)=|\theta|$在$\theta=0$处不可导矛盾，

故$g(\theta)=|\theta|$没有无偏估计.

11．设总体$X$服从正态分布$N(\mu,\sigma^2)$，$X_1, X_2, \cdots, X_n$为来自总体$X$的样本，为了得到标准差$\sigma$的估计量，考虑统计量：

$$Y_1=\frac{1}{n}\sum_{i=1}^{n}|X_i-\overline{X}|,\quad \overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad n\ge 2,$$

$$Y_2=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}|X_i-X_j|,\quad n\ge 2,$$

求常数$C_1$与$C_2$，使得$C_1Y_1$与$C_2Y_2$都是$\sigma$的无偏估计.

解：设$Y\sim N(0,\theta^2)$，有

$$E[|Y|]=\int_{-\infty}^{+\infty}|y|\cdot\frac{1}{\sqrt{2\pi}\theta}\mathrm{e}^{-\frac{y^2}{2\cdot\theta^2}}dy=2\int_0^{+\infty}y\cdot\frac{1}{\sqrt{2\pi}\theta}\mathrm{e}^{-\frac{y^2}{2\theta^2}}dy=-2\frac{\theta}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2\theta^2}}\Bigg|_0^{+\infty}=\sqrt{\frac{2}{\pi}}\theta,$$

因$X_i-\overline{X}$是独立正态变量$X_1, X_2, \cdots, X_n$的线性组合，

且$E(X_i-\overline{X})=E(X_i)-E(\overline{X})=\mu-\mu=0$，

$$\mathrm{Var}(X_i-\overline{X})=\mathrm{Var}(X_i)+\mathrm{Var}(\overline{X})-2\,\mathrm{Cov}(X_i,\overline{X})=\sigma^2+\frac{1}{n}\sigma^2-2\,\mathrm{Cov}\left(X_i,\frac{1}{n}X_i\right)=\frac{n-1}{n}\sigma^2,$$

则$X_i-\overline{X}\sim N\left(0,\dfrac{n-1}{n}\sigma^2\right)$，$E[|X_i-\overline{X}|]=\sqrt{\dfrac{2}{\pi}}\cdot\sqrt{\dfrac{n-1}{n}}\sigma=\sqrt{\dfrac{2(n-1)}{n\pi}}\sigma$，

可得$E(C_1Y_1)=C_1E(Y_1)=C_1\cdot\dfrac{1}{n}\sum\limits_{i=1}^{n}E[|X_i-\overline{X}|]=C_1\cdot\dfrac{1}{n}\cdot n\cdot\sqrt{\dfrac{2(n-1)}{n\pi}}\sigma=C_1\sqrt{\dfrac{2(n-1)}{n\pi}}\sigma$，

故当$C_1=\sqrt{\dfrac{n\pi}{2(n-1)}}$时，$E[C_1Y_1]=\sigma$，$C_1Y_1$是$\sigma$的无偏估计；

当$i\ne j$时，$X_i$与$X_j$相互独立，都服从正态分布$N(\mu,\sigma^2)$，
有$E(X_i-X_j)=E(X_i)-E(X_j)=\mu-\mu=0$，$\mathrm{Var}(X_i-X_j)=\mathrm{Var}(X_i)+\mathrm{Var}(X_j)=\sigma^2+\sigma^2=2\sigma^2$，

则 $X_i - X_j \sim N(0, 2\sigma^2)$，$E[|X_i - X_j|] = \sqrt{\dfrac{2}{\pi}} \cdot \sqrt{2}\sigma = \dfrac{2}{\sqrt{\pi}}\sigma$，

当 $i = j$ 时，$X_i - X_j = 0$，$E[|X_i - X_j|] = 0$，

可得 $E(C_2Y_2) = C_2E(Y_2) = C_2 \cdot \dfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}E[|X_i - X_j|] = C_2 \cdot \dfrac{1}{n(n-1)} \cdot (n^2 - n)\dfrac{2}{\sqrt{\pi}}\sigma = C_2\dfrac{2}{\sqrt{\pi}}\sigma$，

故当 $C_2 = \dfrac{\sqrt{\pi}}{2}$ 时，$E[C_2Y_2] = \sigma$，$C_2Y_2$ 是 $\sigma$ 的无偏估计.

# 习题 6.2

1．从一批电子元件中抽取 8 个进行寿命测试，得到如下数据（单位：h）：

1050，1100，1130，1040，1250，1300，1200，1080，

试对这批元件的平均寿命以及寿命分布的标准差给出矩估计.

解：平均寿命 $\mu$ 的矩估计 $\hat{\mu} = \bar{x} = 1143.75$；标准差 $\sigma$ 的矩估计 $\hat{\mu} = s^* = 89.8523$.

2．设总体 $X \sim U(0, \theta)$，现从该总体中抽取容量为 10 的样本，样本值为：

0.5，1.3，0.6，1.7，2.2，1.2，0.8，1.5，2.0，1.6，

试对参数 $\theta$ 给出矩估计.

解：因 $X \sim U(0, \theta)$，有 $E(X) = \dfrac{\theta}{2}$，即 $\theta = 2E(X)$，故 $\theta$ 的矩估计 $\hat{\theta} = 2\bar{x} = 2 \times 1.34 = 2.68$.

3．设总体分布列如下，$X_1, \cdots, X_n$ 是样本，试求未知参数的矩估计.

（1）$P\{X = k\} = \dfrac{1}{N}$，$k = 0, 1, 2, \cdots, N-1$，$N$（正整数）是未知参数；

（2）$P\{X = k\} = (k-1)\theta^2(1-\theta)^{k-2}$，$k = 2, 3, \cdots$，$0 < \theta < 1$.

解：（1）因 $E(X) = \dfrac{1}{N}[0 + 1 + \cdots + (N-1)] = \dfrac{N-1}{2}$，即 $N = 2E(X) + 1$，故 $N$ 的矩估计 $\hat{N} = 2\overline{X} + 1$；

（2）因 $E(X) = \sum_{k=2}^{+\infty} k \cdot (k-1)\theta^2(1-\theta)^{k-2} = \theta^2 \sum_{k=2}^{+\infty}\dfrac{d^2}{d\theta^2}(1-\theta)^k = \theta^2\dfrac{d^2}{d\theta^2}\left[\sum_{k=2}^{+\infty}(1-\theta)^k\right]$

$$= \theta^2\frac{d^2}{d\theta^2}\left[\frac{(1-\theta)^2}{1-(1-\theta)}\right] = \theta^2\frac{d^2}{d\theta^2}\left(\frac{1}{\theta} - 2 + \theta\right) = \theta^2 \cdot \frac{2}{\theta^3} = \frac{2}{\theta},$$

则 $\theta = \dfrac{2}{E(X)}$，

故 $\theta$ 的矩估计 $\hat{\theta} = \dfrac{2}{\overline{X}}$.

4．设总体密度函数如下，$X_1, \cdots, X_n$ 是样本，试求未知参数的矩估计.

（1）$p(x;\theta) = \dfrac{2}{\theta^2}(\theta - x)$，$0 < x < \theta$，$\theta > 0$；

（2）$p(x;\theta) = (\theta + 1)x^\theta$，$0 < x < 1$，$\theta > 0$；

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

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

解：（1）因$E(X)=\int_0^\theta x\cdot\frac{2}{\theta^2}(\theta-x)dx=\frac{2}{\theta^2}\left(\theta\cdot\frac{x^2}{2}-\frac{x^3}{3}\right)\Big|_0^\theta=\frac{\theta}{3}$，即$\theta=3E(X)$，故$\theta$的矩估计$\hat{\theta}=3\overline{X}$；

（2）因$E(X)=\int_0^1 x\cdot(\theta+1)x^\theta dx=(\theta+1)\cdot\frac{x^{\theta+2}}{\theta+2}\Big|_0^1=\frac{\theta+1}{\theta+2}$，即$\theta=\frac{2E(X)-1}{1-E(X)}$，

故$\theta$的矩估计$\hat{\theta}=\frac{2\overline{X}-1}{1-\overline{X}}$；

（3）因$E(X)=\int_0^1 x\cdot\sqrt{\theta}x^{\sqrt{\theta}-1}dx=\sqrt{\theta}\cdot\frac{x^{\sqrt{\theta}+1}}{\sqrt{\theta}+1}\Big|_0^1=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$，即$\theta=\left[\frac{E(X)}{1-E(X)}\right]^2$，

故$\theta$的矩估计$\hat{\theta}=\left(\frac{\overline{X}}{1-\overline{X}}\right)^2$；

（4）因$E(X)=\int_\mu^{+\infty}x\cdot\frac{1}{\theta}\mathrm{e}^{-\frac{x-\mu}{\theta}}dx=\int_\mu^{+\infty}x\cdot(-1)d\,\mathrm{e}^{-\frac{x-\mu}{\theta}}=-x\,\mathrm{e}^{-\frac{x-\mu}{\theta}}\Big|_\mu^{+\infty}+\int_\mu^{+\infty}\mathrm{e}^{-\frac{x-\mu}{\theta}}dx=\mu-\theta\mathrm{e}^{-\frac{x-\mu}{\theta}}\Big|_\mu^{+\infty}=\mu+\theta$，

$$E(X^2)=\int_\mu^{+\infty}x^2\cdot\frac{1}{\theta}\mathrm{e}^{-\frac{x-\mu}{\theta}}dx=\int_\mu^{+\infty}x^2\cdot(-1)d\,\mathrm{e}^{-\frac{x-\mu}{\theta}}=-x^2\,\mathrm{e}^{-\frac{x-\mu}{\theta}}\Big|_\mu^{+\infty}+\int_\mu^{+\infty}2x\,\mathrm{e}^{-\frac{x-\mu}{\theta}}dx=\mu^2+2\theta E(X)$$
$$=\mu^2+2\mu\theta+2\theta^2,$$

则$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\theta^2$，即$\theta=\sqrt{\mathrm{Var}(X)}$，$\mu=E(X)-\sqrt{\mathrm{Var}(X)}$，

故$\theta$的矩估计$\hat{\theta}=S^*$，$\hat{\mu}=\overline{X}-S^*$.

5. 设总体为$N(\mu,1)$，现对该总体观测$n$次，发现有$k$次观测值为正，使用频率替换方法求$\mu$的估计.

解：因$p=P\{X>0\}=P\{X-\mu>-\mu\}=1-\Phi(-\mu)=\Phi(\mu)$，即$\mu=\Phi^{-1}(p)$，

故$\mu$的矩估计$\hat{\mu}=\Phi^{-1}(\hat{p})=\Phi^{-1}\left(\frac{k}{n}\right)$.

6. 甲、乙两个校对员彼此独立对同一本书的样稿进行校对，校完后，甲发现$a$个错字，乙发现$b$个错字，其中共同发现的错字有$c$个，试用矩法给出如下两个未知参数的估计：

（1）该书样稿的总错字个数；

（2）未被发现的错字数.

解：（1）设$N$为该书样稿总错别字个数，且$A$、$B$分别表示甲、乙发现错别字，有$A$与$B$相互独立，

则$P(AB)=P(A)P(B)$，使用频率替换方法，即$\hat{p}_{AB}=\frac{c}{N}=\hat{p}_A\hat{p}_B=\frac{a}{N}\cdot\frac{b}{N}$，得$N=\frac{ab}{c}$，

故总错字个数$N$的矩估计$\hat{N}=\frac{ab}{c}$；

（2）设$k$为未被发现的错字数，因$P(\overline{A}\,\overline{B})=1-P(A\cup B)=1-P(A)-P(B)+P(AB)$，

使用频率替换方法，即$\hat{p}_{\overline{A}\overline{B}}=\frac{k}{N}=1-\hat{p}_A-\hat{p}_B+\hat{p}_{AB}=1-\frac{a}{N}-\frac{b}{N}+\frac{c}{N}$，即$k=N-a-b+c$，

故未被发现的错字数$k$的矩估计$\hat{k}=\hat{N}-a-b+c=\frac{ab}{c}-a-b+c$.

7. 设总体$X$服从二项分布$b(m,p)$，其中$m,p$为未知参数，$X_1,\cdots,X_n$为$X$的一个样本，求$m$与$p$的矩估

计.

解：因 $E(X)=mp$，$\text{Var}(X)=mp(1-p)$，有 $1-p=\dfrac{\text{Var}(X)}{E(X)}$，

则 $p=1-\dfrac{\text{Var}(X)}{E(X)}$，$m=\dfrac{E(X)}{p}=\dfrac{[E(X)]^2}{E(X)-\text{Var}(X)}$，

故 $m$ 的矩估计 $\hat{m}=\dfrac{\overline{X}^2}{\overline{X}-S^{*2}}$，$p$ 的矩估计 $\hat{p}=1-\dfrac{S^{*2}}{\overline{X}}$.

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

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

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

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

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

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

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

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

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

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

且显然$\mu$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$\mu$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 $\hat{\theta}=\overline{X}-\hat{\mu}=\overline{X}-X_{(1)}$；

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

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

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

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

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

解：（1）因 $L(\theta)=\prod_{i=1}^{n}\frac{1}{2\theta}\mathrm{e}^{-|x_i|/\theta}=\frac{1}{2^n\theta^n}\mathrm{e}^{-\frac{1}{\theta}\sum_{i=1}^{n}|x_i|}$，有 $\ln L(\theta)=-n\ln2-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}|x_i|$，

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

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

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

显然$\theta_1$越大且$\theta_2$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$\theta_1$的最大似然估计$\hat{\theta}_1 = X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$，

$\theta_2$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

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

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$\hat{p} = \dfrac{1}{1000} \times 499 = 0.499$．

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

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

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 $p=2-\dfrac{2n}{\sum\limits_{i=1}^{n}x_i}=2-\dfrac{2}{\bar{x}}$，

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

$\sigma^2$ 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 $\theta$ 的无偏估计和相合估计；

（2）求 $\theta$ 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 $\theta$ 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 $\theta$ 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 $\theta$ 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 $\theta$ 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 $p(x)=\begin{cases}\dfrac{1}{\theta}, & \theta<x<2\theta;\\ 0, & 其他.\end{cases}$，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 $p_n(x)=n[F(x)]^{n-1}p(x)=\begin{cases}\dfrac{n(x-\theta)^{n-1}}{\theta^n}, & \theta<x<2\theta;\\ 0, & 其他.\end{cases}$

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

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

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

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

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

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

故 $\theta$ 的最大似然估计 $\hat\theta^*=\dfrac{1}{2}X_{(n)}$ 是参数 $\theta$ 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 $\theta$ 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 $\theta$ 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 $L(\theta)=\prod\limits_{i=1}^n \mathrm{e}^{-(x_i-\theta)}\mathrm{I}_{x_i>\theta}=\mathrm{e}^{-\sum\limits_{i=1}^n x_i+n\theta}\mathrm{I}_{x_1,x_2,\cdots,x_n>\theta}$，

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

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

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

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

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

故 $\hat{\theta}_1=X_{(1)}$ 是 $\theta$ 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 $\theta$ 的矩估计 $\hat{\theta}_2=\overline{X}-1$；

因 $E(X)=\theta+1$，$\mathrm{Var}(X)=\mathrm{Var}(X-\theta)=\theta^2$，

则 $E(\hat{\theta}_2)=E(\overline{X})-1=E(X)-1=\theta$， $\mathrm{Var}(\hat{\theta}_2)=\mathrm{Var}(\overline{X})=\dfrac{1}{n}\mathrm{Var}(X)=\dfrac{\theta^2}{n}$，

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

因 $\lim\limits_{n\to\infty}E(\hat{\theta}_2)=\theta$， $\lim\limits_{n\to\infty}\mathrm{Var}(\hat{\theta}_2)=\lim\limits_{n\to\infty}\dfrac{\theta^2}{n}=0$，

故 $\hat{\theta}_2=\overline{X}-1$ 是 $\theta$ 的相合估计.

9．设总体 $X\sim Exp(1/\theta)$，$X_1, \cdots, X_n$ 是样本，$\theta$ 的矩估计和最大似然估计都是 $\overline{X}$，它也是 $\theta$ 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 $\overline{X}$ 的估计（提示：考虑 $\hat{\theta}_a=a\overline{X}$，找均方误差最小者）.

证：因 $X\sim Exp(1/\theta)$，有 $E(X)=\theta$，$\mathrm{Var}(X)=\theta^2$，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 $\theta$ 的矩估计为 $\hat{\theta}=\overline{X}$；

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 $\theta$ 的最大似然估计也为 $\hat{\theta}=\overline{X}$；

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

故 $\overline{X}$ 是 $\theta$ 的无偏估计；

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

故 $\overline{X}$ 是 $\theta$ 的相合估计；

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

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

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

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

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

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 $\hat{p}=\overline{X}$，

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

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 $\mu,\sigma^2$ 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 $\sigma$，当 $\mu=x$ 时，$L(\mu)$ 取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$，

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

故 $\mu,\sigma^2$ 的最大似然估计不存在.

## 习题 6.4

1．设总体概率函数是 $p(x;\theta)$，$X_1,\cdots,X_n$ 是其样本，$T=T(X_1,\cdots,X_n)$ 是 $\theta$ 的充分统计量，则对 $g(\theta)$ 的任一

估计 $\hat{g}$，令 $\tilde{g}=E(\hat{g}\,|\,T)$，证明：$\mathrm{MSE}(\tilde{g})\le \mathrm{MSE}(\hat{g})$．这说明，在均方误差准则下，人们只需要考虑基于充分估计量的估计．

解：因 $\tilde{g}=E(\hat{g}\,|\,T)$，由 Rao-Blackwell 定理知 $E(\tilde{g})=E(\hat{g})$，$\mathrm{Var}(\tilde{g})\le \mathrm{Var}(\hat{g})$，

故 $\mathrm{MSE}(\tilde{g})=\mathrm{Var}(\tilde{g})+[E(\tilde{g})-g(\theta)]^2\le \mathrm{Var}(\hat{g})+[E(\hat{g})-g(\theta)]^2=\mathrm{MSE}(\hat{g})$．

2．设 $T_1, T_2$ 分别是 $\theta_1, \theta_2$ 的 UMVUE，证明：对任意的（非零）常数 $a, b, aT_1+bT_2$ 是 $a\theta_1+b\theta_2$ 的 UMVUE．

证：因 $T_1, T_2$ 分别是 $\theta_1, \theta_2$ 的 UMVUE，
有 $E(T_1)=\theta_1$，$E(T_2)=\theta_2$，且对任意的满足 $E(\varphi)=0$ 的 $\varphi$ 都有 $\mathrm{Cov}(T_1, \varphi)=\mathrm{Cov}(T_2, \varphi)=0$，
则 $E(aT_1+bT_2)=a\,E(T_1)+b\,E(T_2)=a\theta_1+b\theta_2$，且 $\mathrm{Cov}(aT_1+bT_2, \varphi)=a\,\mathrm{Cov}(T_1, \varphi)+b\,\mathrm{Cov}(T_2, \varphi)=0$，
故 $aT_1+bT_2$ 是 $a\theta_1+b\theta_2$ 的 UMVUE．

3．设 $T$ 是 $g(\theta)$ 的 UMVUE，$\hat{g}$ 是 $g(\theta)$ 的无偏估计，证明，若 $\mathrm{Var}(\hat{g})<+\infty$，则 $\mathrm{Cov}(T, \hat{g})\ge 0$．

证：因 $\hat{g}$ 和 $T$ 都是 $g(\theta)$ 的无偏估计，有 $E(\hat{g})=E(T)=g(\theta)$，即 $E(\hat{g}-T)=0$，

又因 $T$ 是 $g(\theta)$ 的 UMVUE，有 $\mathrm{Cov}(T, \hat{g}-T)=0$，即 $\mathrm{Cov}(T, \hat{g})-\mathrm{Cov}(T,T)=0$，

故 $\mathrm{Cov}(T, \hat{g})=\mathrm{Cov}(T,T)\ge 0$．

4．设总体 $X\sim N(\mu, \sigma^2)$，$X_1, \cdots, X_n$ 为样本，证明，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 分别为 $\mu, \sigma^2$ 的 UMVUE．

证：因 $X\sim N(\mu, \sigma^2)$，有 $\overline{X}$ 是 $\mu$ 的无偏估计，$S^2$ 是 $\sigma^2$ 的无偏估计，且样本 $X_1, \cdots, X_n$ 的联合密度函数为

$$p(x_1,\cdots,x_n;\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2},$$

对任意的满足 $E(\varphi)=0$ 的 $\varphi(x_1, \cdots, x_n)$，有 $E(\varphi)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n=0$，

对 $E(\varphi)=0$ 两端关于 $\mu$ 求偏导数，得

$$\frac{\partial E(\varphi)}{\partial\mu}=0=\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\frac{1}{\sigma^2}(n\bar{x}-n\mu)\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\frac{n}{\sigma^2}E[(\overline{X}-\mu)\varphi]=\frac{n}{\sigma^2}[E(\overline{X}\varphi)-\mu\,E(\varphi)]=\frac{n}{\sigma^2}E(\overline{X}\varphi),$$

则 $E(\overline{X}\varphi)=0$，$\mathrm{Cov}(\overline{X}, \varphi)=E(\overline{X}\varphi)-E(\overline{X})\cdot E(\varphi)=0$，

故 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是 $\mu$ 的 UMVUE；

对 $E(\overline{X}\varphi)=0$ 两端再关于 $\mu$ 求偏导数，得

$$\frac{\partial E(\overline{X}\varphi)}{\partial\mu}=0=\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\overline{x}\varphi\cdot\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\overline{x}\varphi\cdot\frac{1}{\sigma^2}(n\overline{x}-n\mu)\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\frac{n}{\sigma^2}E[(\overline{X}-\mu)\overline{X}\varphi]=\frac{n}{\sigma^2}[E(\overline{X}^2\varphi)-\mu E(\overline{X}\varphi)]=\frac{n}{\sigma^2}E(\overline{X}^2\varphi),$$

则 $E(\overline{X}^2\varphi)=0$，

对 $(\sqrt{2\pi}\sigma)^n E(\varphi)=0$ 两端关于 $\sigma^2$ 求偏导数，得

$$\frac{\partial[(\sqrt{2\pi}\sigma)^n E(\varphi)]}{\partial\sigma^2}=0=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\frac{1}{2\sigma^4}\left(\sum_{i=1}^{n}x_i^2-2n\overline{x}\mu+n\mu^2\right)\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}dx_1\cdots dx_n$$

$$=\frac{(\sqrt{2\pi}\sigma)^n}{2\sigma^4}E\left[\left(\sum_{i=1}^{n}X_i^2-2n\overline{X}\mu+n\mu^2\right)\varphi\right]$$

$$=\frac{(\sqrt{2\pi}\sigma)^n}{2\sigma^4}\left[E\left(\varphi\sum_{i=1}^{n}X_i^2\right)-2n\mu E(\overline{X}\varphi)+n\mu^2E(\varphi)\right]=\frac{(\sqrt{2\pi}\sigma)^n}{2\sigma^4}E\left(\varphi\sum_{i=1}^{n}X_i^2\right),$$

则 $E\left(\varphi\sum_{i=1}^{n}X_i^2\right)=0$，

因 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$，有 $E(S^2\varphi)=\frac{1}{n-1}\left[E\left(\varphi\sum_{i=1}^{n}X_i^2\right)-nE(\overline{X}^2\varphi)\right]=0$，

则 $\operatorname{Cov}(S^2,\varphi)=E(S^2\varphi)-E(S^2)\cdot E(\varphi)=0$，

故 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是 $\sigma^2$ 的 UMVUE.

5． 设总体的概率函数为 $p(x;\theta)$，满足定义 6.4.2 的条件，若二阶导数 $\frac{\partial^2}{\partial\theta^2}p(x;\theta)$ 对一切的 $\theta\in\Theta$ 存在，证明费希尔信息量 $I(\theta)=-E\left(\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right)$.

证：因 $\frac{\partial}{\partial\theta}\ln p=\frac{1}{p}\cdot\frac{\partial p}{\partial\theta}$，$\frac{\partial^2}{\partial\theta^2}\ln p=\frac{\partial}{\partial\theta}\left(\frac{1}{p}\cdot\frac{\partial p}{\partial\theta}\right)=-\frac{1}{p^2}\cdot\left(\frac{\partial p}{\partial\theta}\right)^2+\frac{1}{p}\cdot\frac{\partial^2 p}{\partial\theta^2}=-\left(\frac{\partial}{\partial\theta}\ln p\right)^2+\frac{1}{p}\cdot\frac{\partial^2 p}{\partial\theta^2}$，

故 $E\left(\frac{\partial^2}{\partial\theta^2}\ln p\right)=-E\left(\frac{\partial}{\partial\theta}\ln p\right)^2+E\left(\frac{1}{p}\cdot\frac{\partial^2 p}{\partial\theta^2}\right)=-I(\theta)+\int_{-\infty}^{+\infty}\frac{1}{p}\cdot\frac{\partial^2 p}{\partial\theta^2}\cdot p\,dx=-I(\theta)+\int_{-\infty}^{+\infty}\frac{\partial^2 p}{\partial\theta^2}dx$

$$= -I(\theta) + \frac{\partial^2}{\partial \theta^2}\left(\int_{-\infty}^{+\infty} p(x)dx\right) = -I(\theta) .$$

6． 设总体密度函数为 $p(x;\theta) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$，$X_1, \cdots, X_n$ 是样本.

（1）求 $g(\theta) = 1/\theta$ 的最大似然估计；

（2）求 $g(\theta)$的有效估计.

解：（1）似然函数 $L(\theta) = \prod_{i=1}^{n} \theta x_i^{\theta-1} \mathrm{I}_{0<x_i<1} = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1} \mathrm{I}_{0<x_1, x_2, \cdots, x_n<1}$，

当 $0 < x_1, x_2, \cdots, x_n < 1$ 时，$\ln L(\theta) = n\ln\theta + (\theta-1)\ln(x_1 x_2 \cdots x_n)$，

令 $\frac{d\ln L(\theta)}{d\theta} = \frac{n}{\theta} + \ln(x_1 x_2 \cdots x_n) = 0$，得 $\theta = -\frac{n}{\ln(x_1 x_2 \cdots x_n)} = -\frac{n}{\sum_{i=1}^{n} \ln x_i}$，即 $\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}$，

故 $g(\theta) = 1/\theta$ 的最大似然估计为 $\hat{g} = 1/\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i$；

（2）因 $E(\ln X) = \int_0^1 \ln x \cdot \theta x^{\theta-1} dx = \int_0^1 \ln x \cdot d(x^\theta) = x^\theta \ln x\Big|_0^1 - \int_0^1 x^\theta \cdot \frac{1}{x} dx = 0 - \frac{1}{\theta} x^\theta \Big|_0^1 = -\frac{1}{\theta}$，

$$E(\ln X)^2 = \int_0^1 (\ln x)^2 \cdot \theta x^{\theta-1} dx = \int_0^1 (\ln x)^2 d(x^\theta) = x^\theta (\ln x)^2 \Big|_0^1 - \int_0^1 x^\theta \cdot \frac{2\ln x}{x} dx = -\frac{2}{\theta} E(\ln X) = \frac{2}{\theta^2},$$

则 $\mathrm{Var}(\ln X) = E(\ln X)^2 - [E(\ln X)]^2 = \frac{2}{\theta^2} - \left(-\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$，

可得 $E(\hat{g}) = -\frac{1}{n}\sum_{i=1}^{n} E(\ln X_i) = -\frac{1}{n} \cdot n \cdot \left(-\frac{1}{\theta}\right) = \frac{1}{\theta} = g(\theta)$，即 $\hat{g} = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ 是 $g(\theta)$的无偏估计，

且 $\mathrm{Var}(\hat{g}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(\ln X_i) = \frac{1}{n^2} \cdot n \cdot \frac{1}{\theta^2} = \frac{1}{n\theta^2}$，

因 $p(x;\theta) = \theta x^{\theta-1} \mathrm{I}_{0<x<1}$，当 $0 < x < 1$ 时，$\ln p(x;\theta) = \ln\theta + (\theta-1)\ln x$，

则 $\frac{\partial}{\partial\theta} \ln p(x;\theta) = \frac{1}{\theta} + \ln x$，$\frac{\partial^2}{\partial\theta^2} \ln p(x;\theta) = -\frac{1}{\theta^2}$，即 $I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \ln p(X;\theta)\right] = \frac{1}{\theta^2}$，

可得 $g(\theta) = 1/\theta$ 无偏估计方差的 C-R 下界为 $\frac{[g'(\theta)]^2}{nI(\theta)} = \frac{\left(-\frac{1}{\theta^2}\right)^2}{n \cdot \frac{1}{\theta^2}} = \frac{1}{n\theta^2} = \mathrm{Var}(\hat{g})$，

故 $\hat{g} = -\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ 是 $g(\theta) = 1/\theta$ 的有效估计.

7． 设总体密度函数为 $p(x;\theta) = \frac{2\theta}{x^3} \mathrm{e}^{-\frac{\theta}{x^2}}, x > 0, \theta > 0$，求$\theta$ 的费希尔信息量 $I(\theta)$.

解：因 $p(x;\theta) = \frac{2\theta}{x^3} \mathrm{e}^{-\frac{\theta}{x^2}} \mathrm{I}_{x>0}$，当 $x > 0$ 时，$\ln p(x;\theta) = \ln 2 + \ln\theta - 3\ln x - \frac{\theta}{x^2}$，

则 $\frac{\partial}{\partial\theta}\ln p(x;\theta)=\frac{1}{\theta}-\frac{1}{x^2}$，$\frac{\partial^2}{\partial\theta^2}\ln p(x;\theta)=-\frac{1}{\theta^2}$，

故 $I(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right]=\frac{1}{\theta^2}$．

8．设总体密度函数为 $p(x;\theta)=\theta c^{\theta}x^{-(\theta+1)}, x>c, c>0$ 已知, $\theta>0$，求 $\theta$ 的费希尔信息量 $I(\theta)$.

解：因 $p(x;\theta)=\theta c^{\theta}x^{-(\theta+1)}\mathrm{I}_{x>c}$，当 $x>c$ 时，$\ln p(x;\theta)=\ln\theta+\theta\ln c-(\theta+1)\ln x$，

则 $\frac{\partial}{\partial\theta}\ln p(x;\theta)=\frac{1}{\theta}+\ln c-\ln x$，$\frac{\partial^2}{\partial\theta^2}\ln p(x;\theta)=-\frac{1}{\theta^2}$，

故 $I(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right]=\frac{1}{\theta^2}$．

9．设总体分布列为 $P\{X=x\}=(x-1)\,\theta^2(1-\theta)^{x-2}, x=2,3,\cdots, 0<\theta<1$，求 $\theta$ 的费希尔信息量 $I(\theta)$.

解：因 $p(x;\theta)=(x-1)\,\theta^2(1-\theta)^{x-2}$，有 $\ln p(x;\theta)=\ln(x-1)+2\ln\theta+(x-2)\ln(1-\theta)$,

则 $\frac{\partial}{\partial\theta}\ln p(x;\theta)=\frac{2}{\theta}-\frac{x-2}{1-\theta}$，$\frac{\partial^2}{\partial\theta^2}\ln p(x;\theta)=-\frac{2}{\theta^2}-\frac{x-2}{(1-\theta)^2}$，

可得 $I(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right]=-E\left(-\frac{2}{\theta^2}-\frac{X-2}{(1-\theta)^2}\right)=\frac{2}{\theta^2}+\frac{1}{(1-\theta)^2}[E(X)-2]$，

因 $E(X)=\sum_{k=2}^{+\infty}k\cdot(k-1)\theta^2(1-\theta)^{k-2}=\theta^2\sum_{k=2}^{+\infty}\frac{d^2}{d\theta^2}(1-\theta)^k=\theta^2\frac{d^2}{d\theta^2}\left[\sum_{k=2}^{+\infty}(1-\theta)^k\right]$

$$=\theta^2\frac{d^2}{d\theta^2}\left[\frac{(1-\theta)^2}{1-(1-\theta)}\right]=\theta^2\frac{d^2}{d\theta^2}\left(\frac{1}{\theta}-2+\theta\right)=\theta^2\cdot\frac{2}{\theta^3}=\frac{2}{\theta},$$

故 $I(\theta)=\frac{2}{\theta^2}+\frac{1}{(1-\theta)^2}\left(\frac{2}{\theta}-2\right)=\frac{2}{\theta^2(1-\theta)}$．

10．设 $X_1,\cdots,X_n$ 是来自 $Ga(\alpha,\lambda)$ 的样本，$\alpha>0$ 已知，试证明，$\overline{X}/\alpha$ 是 $g(\lambda)=1/\lambda$ 的有效估计，从而也是 UMVUE.

证：因总体 $X\sim Ga(\alpha,\lambda)$，有 $E(X)=\frac{\alpha}{\lambda}$，$\mathrm{Var}(X)=\frac{\alpha}{\lambda^2}$，

则 $E\left(\frac{\overline{X}}{\alpha}\right)=\frac{1}{\alpha}E(\overline{X})=\frac{1}{\alpha}E(X)=\frac{1}{\alpha}\cdot\frac{\alpha}{\lambda}=\frac{1}{\lambda}=g(\lambda)$，即 $\frac{\overline{X}}{\alpha}$ 是 $g(\lambda)=\frac{1}{\lambda}$ 的无偏估计，

且 $\mathrm{Var}\left(\frac{\overline{X}}{\alpha}\right)=\frac{1}{\alpha^2}\mathrm{Var}(\overline{X})=\frac{1}{\alpha^2}\cdot\frac{1}{n}\mathrm{Var}(X)=\frac{1}{\alpha^2}\cdot\frac{1}{n}\cdot\frac{\alpha}{\lambda^2}=\frac{1}{n\alpha\lambda^2}$，

因 $p(x;\lambda)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}\mathrm{I}_{x>0}$，当 $x>0$ 时，$\ln p(x;\lambda)=\alpha\ln\lambda-\ln\Gamma(\alpha)+(\alpha-1)\ln x-\lambda x$,

则 $\frac{\partial}{\partial\lambda}\ln p(x;\lambda)=\frac{\alpha}{\lambda}-x$，$\frac{\partial^2}{\partial\lambda^2}\ln p(x;\lambda)=-\frac{\alpha}{\lambda^2}$，即 $I(\lambda)=-E\left[\frac{\partial^2}{\partial\lambda^2}\ln p(X;\lambda)\right]=\frac{\alpha}{\lambda^2}$，

可得 $g(\lambda)=1/\lambda$ 无偏估计方差的 C-R 下界为 $\dfrac{[g'(\lambda)]^2}{nI(\lambda)}=\dfrac{\left(-\dfrac{1}{\lambda^2}\right)^2}{n\cdot\dfrac{\alpha}{\lambda^2}}=\dfrac{1}{n\alpha\lambda^2}=\mathrm{Var}\left(\dfrac{\overline{X}}{\alpha}\right)$，

故 $\dfrac{\overline{X}}{\alpha}$ 是 $g(\lambda)=\dfrac{1}{\lambda}$ 的有效估计，从而也是 UMVUE．

11．设 $X_1,\cdots,X_m$ i.i.d. $\sim N(a,\sigma^2)$，$Y_1,\cdots,Y_n$ i.i.d. $\sim N(a,2\sigma^2)$，求 $a$ 和 $\sigma^2$ 的 UMVUE．

解：根据充分性原则，UMVUE 必为充分统计量，先求参数$(a,\sigma^2)$的充分统计量

因样本 $X_1,\cdots,X_m,Y_1,\cdots,Y_n$ 的联合密度函数为

$$p(x_1,\cdots,x_m,y_1,\cdots,y_n;a,\sigma^2)=\prod_{i=1}^m\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x_i-a)^2}{2\sigma^2}}\cdot\prod_{j=1}^n\frac{1}{\sqrt{2\pi}\cdot\sqrt{2}\sigma}\mathrm{e}^{-\frac{(y_j-a)^2}{4\sigma^2}}$$

$$=\frac{1}{(\sqrt{2})^{m+2n}\cdot(\sqrt{\pi}\sigma)^{m+n}}\mathrm{e}^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^m(x_i-a)^2+\frac{1}{2}\sum_{j=1}^n(y_j-a)^2\right]}$$

$$=\frac{1}{(\sqrt{2})^{m+2n}\cdot(\sqrt{\pi}\sigma)^{m+n}}\mathrm{e}^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^m x_i^2+\frac{1}{2}\sum_{j=1}^n y_j^2-2a\left(\sum_{i=1}^m x_i+\frac{1}{2}\sum_{j=1}^n y_j\right)+\left(m+\frac{n}{2}\right)a^2\right]},$$

令 $(T_1,T_2)=\left(\sum_{i=1}^m X_i+\dfrac{1}{2}\sum_{j=1}^n Y_j,\sum_{i=1}^m X_i^2+\dfrac{1}{2}\sum_{j=1}^n Y_j^2\right)$，有 $(t_1,t_2)=\left(\sum_{i=1}^m x_i+\dfrac{1}{2}\sum_{j=1}^n y_j,\sum_{i=1}^m x_i^2+\dfrac{1}{2}\sum_{j=1}^n y_j^2\right)$，

则 $p(x_1,\cdots,x_m,y_1,\cdots,y_n;a,\sigma^2)=\dfrac{1}{(\sqrt{2})^{m+2n}\cdot(\sqrt{\pi}\sigma)^{m+n}}\mathrm{e}^{-\frac{1}{2\sigma^2}[t_2-2at_1+(m+0.5n)a^2]}$，

取 $g(t_1,t_2;a,\sigma^2)=\dfrac{1}{(\sqrt{2})^{m+2n}(\sqrt{\pi}\sigma)^{m+n}}\mathrm{e}^{-\frac{1}{2\sigma^2}[t_2-2at_1+(m+0.5n)a^2]}$，$h(x_1,\cdots,x_m,y_1,\cdots,y_n)=1$ 与参数 $a,\sigma^2$ 无关，

可得 $(T_1,T_2)=\left(\sum_{i=1}^m X_i+\dfrac{1}{2}\sum_{j=1}^n Y_j,\sum_{i=1}^m X_i^2+\dfrac{1}{2}\sum_{j=1}^n Y_j^2\right)$是参数$(a,\sigma^2)$的充分统计量；

因 $E(T_1)=\sum_{i=1}^m E(X_i)+\dfrac{1}{2}\sum_{j=1}^n E(Y_j)=(m+0.5n)a$，有 $E\left(\dfrac{T_1}{m+0.5n}\right)=E\left(\dfrac{m\overline{X}+0.5n\overline{Y}}{m+0.5n}\right)=a$，

则 $\hat{a}=\dfrac{m\overline{X}+0.5n\overline{Y}}{m+0.5n}$ 是参数 $a$ 的无偏估计，

对任意的满足 $E(\varphi)=0$ 的统计量 $\varphi(x_1,\cdots,x_m,y_1,\cdots,y_n)$，

有 $E(\varphi)=\dfrac{1}{(\sqrt{2})^{m+2n}\cdot(\sqrt{\pi}\sigma)^{m+n}}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(m+0.5n)a^2}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

则 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

两端关于 $a$ 求偏导数，得 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}\cdot\dfrac{1}{2\sigma^2}\cdot 2t_1dx_1\cdots dx_m dy_1\cdots dy_n=0$，

即 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}t_1\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

则 $E(T_1\varphi)=0$，有 $E(\hat{a}\varphi)=\dfrac{1}{m+0.5n}E(T_1\varphi)=0$，即 $\mathrm{Cov}(\hat{a},\varphi)=E(\hat{a}\varphi)-E(\hat{a})E(\varphi)=0$，

故 $\hat{a}=\dfrac{m\overline{X}+0.5n\overline{Y}}{m+0.5n}$ 是参数 $a$ 的 UMVUE；

因 $E(T_2)=\sum\limits_{i=1}^{m}E(X_i^2)+\dfrac{1}{2}\sum\limits_{j=1}^{n}E(Y_j^2)=\sum\limits_{i=1}^{m}(\sigma^2+a^2)+\dfrac{1}{2}\sum\limits_{j=1}^{n}(2\sigma^2+a^2)=(m+n)\sigma^2+(m+0.5n)a^2$，

且 $E(T_1^2)=\mathrm{Var}(T_1)+[E(T_1)]^2=\sum\limits_{i=1}^{m}\mathrm{Var}(X_i)+\dfrac{1}{4}\sum\limits_{j=1}^{n}\mathrm{Var}(Y_j)+[(m+0.5n)a]^2$

$=(m+0.5n)\sigma^2+(m+0.5n)^2a^2$，

则 $E\left(T_2-\dfrac{T_1^2}{m+0.5n}\right)=(m+n-1)\sigma^2$，即 $E\left[\dfrac{1}{m+n-1}\left(T_2-\dfrac{T_1^2}{m+0.5n}\right)\right]=\sigma^2$，

取 $\hat{\sigma}^2=\dfrac{1}{m+n-1}\left(T_2-\dfrac{T_1^2}{m+0.5n}\right)=\dfrac{1}{m+n-1}\left[\sum\limits_{i=1}^{m}X_i^2+\dfrac{1}{2}\sum\limits_{j=1}^{n}Y_j^2-\dfrac{1}{m+0.5n}\left(\sum\limits_{i=1}^{m}X_i+\dfrac{1}{2}\sum\limits_{j=1}^{n}Y_j\right)^2\right]$，

可知 $\hat{\sigma}^2$ 是参数 $\sigma^2$ 的无偏估计，

因 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

两端关于 $\sigma^2$ 求偏导数，得 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}\cdot\dfrac{1}{2\sigma^4}\cdot(t_2-2at_1)dx_1\cdots dx_m dy_1\cdots dy_n=0$，

即 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}(t_2-2at_1)\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$

则 $E[(T_2-2aT_1)\varphi]=0$，有 $E(T_2\varphi)-2a\,E(T_1\varphi)=0$，可得 $E(T_2\varphi)=0$，

又因 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}t_1\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

两端关于 $a$ 求偏导数，得 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}t_1\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}\cdot\dfrac{1}{2\sigma^2}\cdot 2at_1dx_1\cdots dx_m dy_1\cdots dy_n=0$，

即 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}t_1^2\varphi\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}(t_2-2at_1)}dx_1\cdots dx_m dy_1\cdots dy_n=0$，

则 $E(T_1^2\varphi)=0$，有 $E(\hat{\sigma}^2\varphi)=\dfrac{1}{m+n-1}\left[E(T_2\varphi)-\dfrac{E(T_1^2\varphi)}{m+0.5n}\right]=0$，

即 $\mathrm{Cov}(\hat{\sigma}^2,\varphi)=E(\hat{\sigma}^2\varphi)-E(\hat{\sigma}^2)E(\varphi)=0$，

故 $\hat{\sigma}^2=\dfrac{1}{m+n-1}\left[\sum\limits_{i=1}^{m}X_i^2+\dfrac{1}{2}\sum\limits_{j=1}^{n}Y_j^2-\dfrac{1}{m+0.5n}\left(\sum\limits_{i=1}^{m}X_i+\dfrac{1}{2}\sum\limits_{j=1}^{n}Y_j\right)^2\right]$ 是参数 $\sigma^2$ 的 UMVUE.

12．设 $X_1,\cdots,X_n$ i.i.d. $\sim N(\mu,1)$，求 $\mu^2$ 的 UMVUE．证明此 UMVUE 达不到 C-R 不等式的下界，即它不是有效估计．

解：根据充分性原则，UMVUE 必为充分统计量，先求参数 $\mu^2$ 的充分统计量，

因样本 $X_1,\cdots,X_n$ 的联合密度函数为

$$p(x_1,\cdots,x_n;\mu)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2}}=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)}=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{-\frac{1}{2}\left(\sum_{i=1}^{n}x_i^2-2\mu\sum_{i=1}^{n}x_i+n\mu^2\right)}$$

$$=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{n\mu\bar{x}-\frac{1}{2}n\mu^2}\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2},$$

令 $T=\overline{X}$，有 $t=\bar{x}$，即 $p(x_1,\cdots,x_n;\mu)=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{n\mu t-\frac{1}{2}n\mu^2}\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2}$，

取 $g(t;\mu)=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{n\mu t-\frac{1}{2}n\mu^2}$，$h(x_1,x_2,\cdots,x_n)=\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2}$ 与参数 $\mu$ 无关，

可得 $T=\overline{X}$ 是参数 $\mu$ 的充分统计量；

因 $E(\overline{X}^2)=\mathrm{Var}(\overline{X})+[E(\overline{X})]^2=\frac{1}{n}\mathrm{Var}(X)+[E(X)]^2=\frac{1}{n}+\mu^2$，即 $E\left(\overline{X}^2-\frac{1}{n}\right)=\mu^2$，

可知 $\hat{\mu}^2=\overline{X}^2-\frac{1}{n}$ 是参数 $\mu^2$ 的无偏估计，

对任意的满足 $E(\varphi)=0$ 的统计量 $\varphi(x_1,\cdots,x_n)$，

有 $E(\varphi)=\frac{1}{(\sqrt{2\pi})^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2+n\mu\bar{x}-\frac{1}{2}n\mu^2}dx_1\cdots dx_n=0$，

则 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2+n\mu\bar{x}}dx_1\cdots dx_n=0$，

两端关于 $\mu$ 求偏导数，得 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2+n\mu\bar{x}}\cdot n\bar{x}dx_1\cdots dx_n=0$，

两端关于 $\mu$ 再求偏导数，得 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\varphi\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2+n\mu\bar{x}}\cdot(n\bar{x})^2dx_1\cdots dx_n=0$，

即 $\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\bar{x}^2\varphi\cdot\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}x_i^2+n\mu\bar{x}}dx_1\cdots dx_n=0$，

则 $E(\overline{X}^2\varphi)=0$，有 $E(\hat{\mu}^2\varphi)=E(\overline{X}^2\varphi)-\frac{1}{n}E(\varphi)=0$，即 $\mathrm{Cov}(\hat{\mu}^2,\varphi)=E(\hat{\mu}^2\varphi)-E(\hat{\mu}^2)E(\varphi)=0$，

故 $\hat{\mu}^2=\overline{X}^2-\frac{1}{n}$ 是参数 $\mu^2$ 的 UMVUE；

因 $\overline{X}\sim N\left(\mu,\frac{1}{n}\right)$，有 $E(\overline{X})=\mu$，$\mathrm{Var}(\overline{X})=E[(\overline{X}-\mu)^2]=\frac{1}{n}$，$E[(\overline{X}-\mu)^3]=0$，$E[(\overline{X}-\mu)^4]=\frac{3}{n^2}$，

则 $E(\overline{X}^4)=E[(\overline{X}-\mu+\mu)^4]=E[(\overline{X}-\mu)^4]+4\mu E[(\overline{X}-\mu)^3]+6\mu^2E[(\overline{X}-\mu)^2]+4\mu^3E(\overline{X}-\mu)+\mu^4$

$$=\frac{3}{n^2}+\frac{6\mu^2}{n}+\mu^4,$$

可得 $\mathrm{Var}(\hat{\mu}^2)=\mathrm{Var}(\overline{X}^2)=E(\overline{X}^4)-[E(\overline{X}^2)]^2=\dfrac{3}{n^2}+\dfrac{6\mu^2}{n}+\mu^4-\left(\dfrac{1}{n}+\mu^2\right)^2=\dfrac{2}{n^2}+\dfrac{4\mu^2}{n}$，

因总体密度函数 $p(x;\mu)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2}}$，有 $\ln p(x;\mu)=-\ln\sqrt{2\pi}-\dfrac{(x-\mu)^2}{2}$，

则 $\dfrac{\partial}{\partial\mu}\ln p(x;\mu)=x-\mu$，即 $I(\mu)=E\left(\dfrac{\partial}{\partial\mu}\ln p(X;\mu)\right)^2=E(X-\mu)^2=1$，

可得 $g(\mu)=\mu^2$ 无偏估计方差的 C-R 下界为 $\dfrac{[g'(\mu)]^2}{nI(\mu)}=\dfrac{(2\mu)^2}{n}=\dfrac{4\mu^2}{n}<\mathrm{Var}(\hat{\mu}^2)$，

故 $\hat{\mu}^2=\overline{X}^2-\dfrac{1}{n}$ 不是参数 $\mu^2$ 的有效估计.

13．对泊松分布 $P(\theta)$.

（1）求 $I\left(\dfrac{1}{\theta}\right)$；

（2）找一个函数 $g(\cdot)$，使 $g(\theta)$ 的费希尔信息与 $\theta$ 无关.

解：因总体概率函数为 $p(x;\alpha)=\dfrac{\theta^x}{x!}\mathrm{e}^{-\theta}$，有 $\ln p(x;\theta)=x\ln\theta-\ln x!-\theta$,

则 $\dfrac{\partial}{\partial\theta}\ln p(x;\theta)=x\cdot\dfrac{1}{\theta}-1=\dfrac{x-\theta}{\theta}$，即 $I(\theta)=E\left[\dfrac{\partial}{\partial\theta}\ln p(X;\theta)\right]^2=\dfrac{1}{\theta^2}E(X-\theta)^2=\dfrac{1}{\theta^2}\mathrm{Var}(X)=\dfrac{1}{\theta}$，

令 $\alpha=g(\theta)$ 可导，有 $\dfrac{\partial}{\partial\theta}\ln p=\dfrac{\partial}{\partial\alpha}\ln p\cdot\dfrac{d\alpha}{d\theta}=g'(\theta)\cdot\dfrac{\partial}{\partial\alpha}\ln p$，

则 $I(\theta)=E\left[\dfrac{\partial}{\partial\theta}\ln p\right]^2=[g'(\theta)]^2E\left[\dfrac{\partial}{\partial\alpha}\ln p\right]^2=[g'(\theta)]^2I(\alpha)=[g'(\theta)]^2I[g(\theta)]$，即 $I[g(\theta)]=\dfrac{I(\theta)}{[g'(\theta)]^2}$，

（1）因 $g(\theta)=\dfrac{1}{\theta}$，有 $g'(\theta)=-\dfrac{1}{\theta^2}$，

故 $I\left(\dfrac{1}{\theta}\right)=\dfrac{I(\theta)}{[g'(\theta)]^2}=\dfrac{1/\theta}{(-1/\theta^2)^2}=\theta^3$；

（2）要使得 $I[g(\theta)]=\dfrac{I(\theta)}{[g'(\theta)]^2}=\dfrac{1}{\theta[g'(\theta)]^2}=c$ 为常数与 $\theta$ 无关，

则 $[g'(\theta)]^2=\dfrac{1}{c\theta}$，$g'(\theta)=\dfrac{1}{\sqrt{c\theta}}$，即 $g(\theta)=\dfrac{2}{\sqrt{c}}\sqrt{\theta}$，

取 $g(\theta)=\sqrt{\theta}$，有 $g'(\theta)=\dfrac{1}{2\sqrt{\theta}}$，

故 $I[g(\theta)]=\dfrac{I(\theta)}{[g'(\theta)]^2}=\dfrac{1/\theta}{[1/(2\sqrt{\theta})]^2}=4$ 与 $\theta$ 无关.

14．设 $X_1,\cdots,X_n$ 为独立同分布变量，$0<\theta<1$，

$$P\{X_1=-1\}=\frac{1-\theta}{2},\quad P\{X_1=0\}=\frac{1}{2},\quad P\{X_1=1\}=\frac{\theta}{2}.$$

（1）求$\theta$的 MLE $\hat{\theta}_1$，并问$\hat{\theta}_1$是否无偏的；

（2）求$\theta$的矩估计$\hat{\theta}_2$；

（3）计算$\theta$的无偏估计的方差的 C-R 下界.

解：（1）方法一：设 $X_1, \cdots, X_n$ 中取值$-1, 0, 1$ 分别有 $n_{-1}, n_0, n_1$ 次，有 $n_{-1} + n_0 + n_1 = n$，

则似然函数 $L(\theta)=\left(\dfrac{1-\theta}{2}\right)^{n_{-1}}\left(\dfrac{1}{2}\right)^{n_0}\left(\dfrac{\theta}{2}\right)^{n_1}=\dfrac{(1-\theta)^{n_{-1}}\theta^{n_1}}{2^n}$，有 $\ln L(\theta) = n_{-1}\ln(1-\theta) + n_1\ln\theta - n\ln 2$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=n_{-1}\cdot\dfrac{-1}{1-\theta}+n_1\cdot\dfrac{1}{\theta}=0$，得 $\theta=\dfrac{n_1}{n_{-1}+n_1}$，

故$\theta$的 MLE $\hat{\theta}_1=\dfrac{n_1}{n_{-1}+n_1}$；

方法二：总体 $X$ 概率函数为

$$p(x;\theta)=\left(\frac{1-\theta}{2}\right)^{\frac{x(x-1)}{2}}\left(\frac{1}{2}\right)^{-(x+1)(x-1)}\left(\frac{\theta}{2}\right)^{\frac{x(x+1)}{2}}=\frac{1}{2}(1-\theta)^{\frac{x^2-x}{2}}\theta^{\frac{x^2+x}{2}}，\ x=-1, 0, 1，$$

则似然函数 $L(\theta)=\displaystyle\prod_{i=1}^n\frac{1}{2}(1-\theta)^{\frac{x_i^2-x_i}{2}}\theta^{\frac{x_i^2+x_i}{2}}=\frac{1}{2^n}(1-\theta)^{\frac{1}{2}\left(\sum_{i=1}^n x_i^2-\sum_{i=1}^n x_i\right)}\theta^{\frac{1}{2}\left(\sum_{i=1}^n x_i^2+\sum_{i=1}^n x_i\right)}$，

有 $\ln L(\theta)=\dfrac{1}{2}\left(\displaystyle\sum_{i=1}^n x_i^2-\sum_{i=1}^n x_i\right)\ln(1-\theta)+\dfrac{1}{2}\left(\displaystyle\sum_{i=1}^n x_i^2+\sum_{i=1}^n x_i\right)\ln\theta-n\ln 2$，

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

故$\theta$的 MLE $\hat{\theta}_1=\dfrac{1}{2}+\dfrac{\sum_{i=1}^n X_i}{2\sum_{i=1}^n X_i^2}$；

（注：因 $X_i$ 全部可能取值$-1, 0, 1$，有$\displaystyle\sum_{i=1}^n X_i^2=n_{-1}+n_1$，$\displaystyle\sum_{i=1}^n X_i=n_1-n_{-1}$，即以上两个结果一致）

因 $E(\hat{\theta}_1)=E\left(\dfrac{n_1}{n_{-1}+n_1}\right)=E\left[E\left(\dfrac{n_1}{n_{-1}+n_1}\middle|n_{-1}+n_1\right)\right]$，

且 $P\{X=1\mid X=-1或X=1\}=\dfrac{P\{X=1\}}{P\{X=-1或X=1\}}=\dfrac{\frac{\theta}{2}}{\frac{1-\theta}{2}+\frac{\theta}{2}}=\theta$，

则在 $n_{-1}+n_1=m$ 的条件下，$n_1$ 服从二项分布 $b(m, \theta)$，$E(n_1\mid n_{-1}+n_1=m)=m\theta$,

可得 $E\left(\dfrac{n_1}{n_{-1}+n_1}\middle|n_{-1}+n_1=m\right)=\dfrac{1}{m}E(n_1|n_{-1}+n_1=m)=\theta$，即 $E\left(\dfrac{n_1}{n_{-1}+n_1}\middle|n_{-1}+n_1\right)=\theta$，

故 $E(\hat{\theta}_1)=E\left[E\left(\dfrac{n_1}{n_{-1}+n_1}\middle|n_{-1}+n_1\right)\right]=E(\theta)=\theta$，$\hat{\theta}_1$ 是 $\theta$ 的无偏估计；

（2）因 $E(X)=(-1)\times\dfrac{1-\theta}{2}+0\times\dfrac{1}{2}+1\times\dfrac{\theta}{2}=\theta-\dfrac{1}{2}$，有 $\theta=E(X)+\dfrac{1}{2}$，

故 $\theta$ 的矩估计 $\hat{\theta}_2=\overline{X}-\dfrac{1}{2}$；

（3）因总体 $X$ 概率函数为 $p(x;\theta)=\dfrac{1}{2}(1-\theta)^{\frac{x^2-x}{2}}\theta^{\frac{x^2+x}{2}}$，$x=-1, 0, 1$，

有 $\ln p(x;\theta)=\dfrac{x^2-x}{2}\ln(1-\theta)+\dfrac{x^2+x}{2}\ln\theta-\ln 2$，

则 $\dfrac{\partial}{\partial\theta}\ln p(x;\theta)=\dfrac{x^2-x}{2}\cdot\dfrac{-1}{1-\theta}+\dfrac{x^2+x}{2}\cdot\dfrac{1}{\theta}$，

即 $\dfrac{\partial^2}{\partial\theta^2}\ln p(x;\theta)=\dfrac{x^2-x}{2}\cdot\dfrac{-1}{(1-\theta)^2}-\dfrac{x^2+x}{2}\cdot\dfrac{1}{\theta^2}=-\dfrac{[(1-\theta)^2+\theta^2]x^2+[(1-\theta)^2-\theta^2]x}{2\theta^2(1-\theta)^2}$，

可得费希尔信息量 $I(\theta)=-E\left[\dfrac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right]=\dfrac{[(1-\theta)^2+\theta^2]E(X^2)+[(1-\theta)^2-\theta^2]E(X)}{2\theta^2(1-\theta)^2}$，

因 $E(X)=(-1)\times\dfrac{1-\theta}{2}+0\times\dfrac{1}{2}+1\times\dfrac{\theta}{2}=\theta-\dfrac{1}{2}$，$E(X^2)=(-1)^2\times\dfrac{1-\theta}{2}+0^2\times\dfrac{1}{2}+1^2\times\dfrac{\theta}{2}=\dfrac{1}{2}$，

则 $I(\theta)=\dfrac{(2\theta^2-2\theta+1)\cdot\dfrac{1}{2}+(1-2\theta)\cdot\left(\theta-\dfrac{1}{2}\right)}{2\theta^2(1-\theta)^2}=\dfrac{\theta-\theta^2}{2\theta^2(1-\theta)^2}=\dfrac{1}{2\theta(1-\theta)}$，

故 $\theta$ 的 C-R 下界为 $\dfrac{1}{nI(\theta)}=\dfrac{2\theta(1-\theta)}{n}$.

15．设总体 $X\sim Exp(1/\theta)$，$X_1, \cdots, X_n$ 是样本，$\theta$ 的矩估计和最大似然估计都是 $\overline{X}$，它也是 $\theta$ 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 $\overline{X}$ 的估计（提示：考虑 $\hat{\theta}_a=a\overline{X}$，找均方误差最小者）.

注：此题与习题 6.3 第 9 题相同，这里省略.

## 习题 6.5

1．设一页书上的错别字个数服从泊松分布 $P(\lambda)$，有两个可能取值：1.5 和 1.8，且先验分布为

$$P\{\lambda=1.5\}=0.45,\quad P\{\lambda=1.8\}=0.55,$$

现检查了一页，发现有 3 个错别字，试求 $\lambda$ 的后验分布.

解：总体 $X$ 表示一页书上的错别字个数，$X\sim P(\lambda)$，样本为 $X_1=3$，有 $P\{X_1=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$，$k=0,1,2,\cdots$，

则 $P\{X_1=3\}=P\{\lambda=1.5\}P\{X_1=3\mid\lambda=1.5\}+P\{\lambda=1.8\}P\{X_1=3\mid\lambda=1.8\}$

$$=0.45\times\frac{1.5^3}{6}\cdot e^{-1.5}+0.55\times\frac{1.8^3}{6}\cdot e^{-1.8}=0.0565+0.0884=0.1449,$$

故参数 $\lambda$ 的后验分布为 $P\{\lambda=1.5\mid X_1=3\}=\dfrac{P\{\lambda=1.5\}P\{X_1=3\mid\lambda=1.5\}}{P\{X_1=3\}}=\dfrac{0.0565}{0.1449}=0.3899$，

$$P\{\lambda=1.8 \mid X_1=3\}=\frac{P\{\lambda=1.8\}P\{X_1=3 \mid \lambda=1.8\}}{P\{X_1=3\}}=\frac{0.0884}{0.1449}=0.6101 .$$

2． 设总体为均匀分布 $U(\theta, \theta+1)$，$\theta$ 的先验分布是均匀分布 $U(10, 16)$. 现有三个观测值：11.7, 12.1, 12.0. 求 $\theta$ 的后验分布.

解：参数 $\theta$ 的先验分布为 $\pi(\theta)=\frac{1}{6}\mathrm{I}_{10<\theta<16}$，

总体 $X$ 的条件分布为 $p(x \mid \theta)=\mathrm{I}_{\theta<x<\theta+1}$，

有样本 $X_1, X_2, X_3$ 的联合条件分布为 $p(x_1, x_2, x_3 \mid \theta)=\mathrm{I}_{\theta<x_1, x_2, x_3<\theta+1}$，

则样本 $X_1, X_2, X_3$ 和参数 $\theta$ 的联合分布为

$$h(x_1, x_2, x_3, \theta)=\frac{1}{6}\mathrm{I}_{\theta<x_1, x_2, x_3<\theta+1, 10<\theta<16}=\frac{1}{6}\mathrm{I}_{x_{(3)}-1<\theta<x_{(1)}, 10<\theta<16},$$

可得样本 $X_1, X_2, X_3$ 的边际分布为 $m(x_1, x_2, x_3)=\int_{-\infty}^{+\infty}\frac{1}{6}\mathrm{I}_{x_{(3)}-1<\theta<x_{(1)}, 10<\theta<16}d\theta=\int_{11.1}^{11.7}\frac{1}{6}d\theta=0.1$，

故参数 $\theta$ 的后验分布为 $\pi(\theta \mid x_1, x_2, x_3)=\frac{h(x_1, x_2, x_3, \theta)}{m(x_1, x_2, x_3)}=\frac{5}{3}\mathrm{I}_{11.1<\theta<11.7}$ .

3． 设 $X_1, \cdots, X_n$ 是来自几何分布的样本，总体分布列为

$$P\{X=k \mid \theta\}=\theta(1-\theta)^k, \quad k=0, 1, 2, \cdots,$$

$\theta$ 的先验分布是均匀分布 $U(0, 1)$.

（1）求 $\theta$ 的后验分布；

（2）若 4 次观测值为 4, 3, 1, 6，求 $\theta$ 的贝叶斯估计.

解：（1）参数 $\theta$ 的先验分布为 $\pi(\theta)=\mathrm{I}_{0<\theta<1}$，

因样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1, \cdots, x_n \mid \theta)=\prod_{i=1}^{n}\theta(1-\theta)^{x_i}=\theta^n(1-\theta)^{x_1+\cdots+x_n}, \quad x_1, \cdots, x_n=0, 1, 2, \cdots,$$

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1, \cdots, x_n, \theta)=\theta^n(1-\theta)^{x_1+\cdots+x_n}\mathrm{I}_{0<\theta<1}, \quad x_1, \cdots, x_n=0, 1, 2, \cdots,$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1, \cdots, x_n)=\int_0^1\theta^n(1-\theta)^{x_1+\cdots+x_n}d\theta=\frac{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}{\Gamma(n+x_1+\cdots+x_n+2)}, \quad x_1, \cdots, x_n=0, 1, 2, \cdots,$$

故参数 $\theta$ 的后验分布为

$$\pi(\theta \mid x_1, \cdots, x_n)=\frac{h(x_1, \cdots, x_n, \theta)}{m(x_1, \cdots, x_n)}=\frac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\theta^n(1-\theta)^{x_1+\cdots+x_n}\mathrm{I}_{0<\theta<1};$$

（2）因 $E(\theta \mid x_1, \cdots, x_n)=\int_0^1\theta\cdot\pi(\theta \mid x_1, \cdots, x_n)d\theta=\frac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\int_0^1\theta^{n+1}(1-\theta)^{x_1+\cdots+x_n}d\theta$

$$=\frac{\Gamma(n+x_1+\cdots+x_n+2)}{\Gamma(n+1)\Gamma(x_1+\cdots+x_n+1)}\cdot\frac{\Gamma(n+2)\Gamma(x_1+\cdots+x_n+1)}{\Gamma(n+x_1+\cdots+x_n+3)}=\frac{n+1}{n+x_1+\cdots+x_n+2},$$

则贝叶斯估计 $\hat{\theta}_B = E(\theta \mid X_1, \cdots, X_n) = \dfrac{n+1}{n + X_1 + \cdots + X_n + 2}$，

因样本观测值为 4, 3, 1, 6，即 $x_1 + \cdots + x_n = 15$，$n = 4$，

故 $\hat{\theta}_B = \dfrac{4+1}{4+14+2} = \dfrac{1}{4}$.

4． 验证：泊松分布的均值$\lambda$ 的共轭先验分布是伽玛分布.

证：设参数$\lambda$ 的先验分布是伽玛分布 $Ga(\alpha, \beta)$，密度函数为 $\pi(\lambda) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \mathrm{e}^{-\beta\lambda} \mathrm{I}_{\lambda>0}$，

因样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1, \cdots, x_n \mid \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \mathrm{e}^{-\lambda} = \frac{\lambda^{x_1+\cdots+x_n}}{x_1! \cdots x_n!} \mathrm{e}^{-n\lambda}，\ x_1, \cdots, x_n = 0, 1, 2, \cdots,$$

则样本 $X_1, \cdots, X_n$ 和参数$\lambda$ 的联合分布为

$$h(x_1, \cdots, x_n, \lambda) = \frac{\beta^\alpha \lambda^{x_1+\cdots+x_n+\alpha-1}}{\Gamma(\alpha) x_1! \cdots x_n!} \mathrm{e}^{-(n+\beta)\lambda} \mathrm{I}_{\lambda>0}，\ x_1, \cdots, x_n = 0, 1, 2, \cdots,$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1, \cdots, x_n) = \int_0^{+\infty} \frac{\beta^\alpha \lambda^{x_1+\cdots+x_n+\alpha-1}}{\Gamma(\alpha) x_1! \cdots x_n!} \mathrm{e}^{-(n+\beta)\lambda} \, d\lambda = \frac{\beta^\alpha}{\Gamma(\alpha) x_1! \cdots x_n!} \int_0^{+\infty} \lambda^{x_1+\cdots+x_n+\alpha-1} \mathrm{e}^{-(n+\beta)\lambda} \, d\lambda$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha) x_1! \cdots x_n!} \cdot \frac{\Gamma(x_1 + \cdots + x_n + \alpha)}{(n+\beta)^{x_1+\cdots+x_n+\alpha}}，\ x_1, \cdots, x_n = 0, 1, 2, \cdots,$$

即参数$\lambda$ 的后验分布为

$$\pi(\lambda \mid x_1, \cdots, x_n) = \frac{h(x_1, \cdots, x_n, \lambda)}{m(x_1, \cdots, x_n)} = \frac{(n+\beta)^{x_1+\cdots+x_n+\alpha}}{\Gamma(x_1 + \cdots + x_n + \alpha)} \lambda^{x_1+\cdots+x_n+\alpha-1} \mathrm{e}^{-(n+\beta)\lambda} \mathrm{I}_{\lambda>0}，$$

后验分布仍为伽玛分布 $Ga(x_1 + \cdots + x_n + \alpha, n + \beta)$，
故伽玛分布是泊松分布的均值$\lambda$ 的共轭先验分布.

5． 验证：正态总体方差（均值已知）的共轭先验分布是倒伽玛分布.

证：设参数$\sigma^2$ 的先验分布是倒伽玛分布 $IGa(\alpha, \lambda)$，密度函数为 $\pi(\sigma^2) = \dfrac{\lambda^\alpha}{\Gamma(\alpha)} \left(\dfrac{1}{\sigma^2}\right)^{\alpha+1} \mathrm{e}^{-\frac{\lambda}{\sigma^2}}$，

又设总体分布为 $N(\mu_0, \sigma^2)$，其中$\mu_0$ 已知，密度函数为 $p(x \mid \sigma^2) = \dfrac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu_0)^2}{2\sigma^2}}$，

有样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1, \cdots, x_n \mid \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi})^n \sigma^n} \mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu_0)^2}，$$

则样本 $X_1, \cdots, X_n$ 和参数$\sigma^2$ 的联合分布为

$$h(x_1, \cdots, x_n, \sigma^2) = \frac{\lambda^\alpha}{(\sqrt{2\pi})^n \Gamma(\alpha)} \cdot \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1} \mathrm{e}^{-\frac{1}{\sigma^2}\left[\lambda + \frac{1}{2}\sum_{i=1}^n (x_i-\mu_0)^2\right]}，$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1}\mathrm{e}^{-\frac{1}{\sigma^2}\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]}d(\sigma^2)$$

$$=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\int_{+\infty}^0 t^{\frac{n}{2}+\alpha+1}\mathrm{e}^{-t\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]}\left(-\frac{1}{t^2}\right)dt$$

$$=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\int_0^{+\infty} t^{\frac{n}{2}+\alpha-1}\mathrm{e}^{-t\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]}dt=\frac{\lambda^\alpha}{(\sqrt{2\pi})^n\Gamma(\alpha)}\cdot\frac{\Gamma\left(\frac{n}{2}+\alpha\right)}{\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]^{\frac{n}{2}+\alpha}},$$

即参数 $\sigma^2$ 的后验分布为

$$\pi(\sigma^2\mid x_1,\cdots,x_n)=\frac{\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]^{\frac{n}{2}+\alpha}}{\Gamma\left(\frac{n}{2}+\alpha\right)}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+\alpha+1}\mathrm{e}^{-\frac{1}{\sigma^2}\left[\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right]},$$

后验分布仍为倒伽玛分布 $IGa\left(\frac{n}{2}+\alpha,\lambda+\frac{1}{2}\sum\limits_{i=1}^n(x_i-\mu_0)^2\right)$，

故倒伽玛分布是参数 $\sigma^2$ 的共轭先验分布.

6．设 $X_1, \cdots, X_n$ 是来自如下总体的一个样本，

$$p(x\mid\theta)=\frac{2x}{\theta^2},\quad 0<x<\theta.$$

（1）若 $\theta$ 的先验分布为均匀分布 $U(0, 1)$，求 $\theta$ 的后验分布；

（2）若 $\theta$ 的先验分布为 $\pi(\theta)=3\theta^2$，$0<\theta<1$，求 $\theta$ 的后验分布.

解：样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\mid\theta)=\prod_{i=1}^n\frac{2x_i}{\theta^2}\mathrm{I}_{0<x_i<\theta}=\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta},$$

（1）因参数 $\theta$ 的先验分布为 $\pi(\theta)=\mathrm{I}_{0<\theta<1}$，

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta<1}=\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{x_1,\cdots,x_n>0,\,x_{(n)}<\theta<1},$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{x_{(n)}}^1\frac{2^n x_1\cdots x_n}{\theta^{2n}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{2^n x_1\cdots x_n}{2n-1}[x_{(n)}^{-(2n-1)}-1]\cdot\mathrm{I}_{x_1,\cdots,x_n>0},$$

故参数 $\theta$ 的后验分布为

$$\pi(\theta\mid x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{2n-1}{\theta^{2n}[x_{(n)}^{-(2n-1)}-1]}\mathrm{I}_{x_{(n)}<\theta<1};$$

（2）因参数 $\theta$ 的先验分布为 $\pi(\theta)=3\theta^2\,\mathrm{I}_{0<\theta<1}$，

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta<1}=\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{x_1,\cdots,x_n>0,\,x_{(n)}<\theta<1},$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_{x_{(n)}}^{1}\frac{3\cdot 2^n x_1\cdots x_n}{\theta^{2n-2}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{3\cdot 2^n x_1\cdots x_n}{2n-3}[x_{(n)}^{-(2n-3)}-1]\cdot\mathrm{I}_{x_1,\cdots,x_n>0}\text{，}$$

故参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{2n-3}{\theta^{2n-2}[x_{(n)}^{-(2n-3)}-1]}\mathrm{I}_{x_{(n)}<\theta<1}\text{．}$$

7． 设 $X_1, \cdots, X_n$ 是来自如下总体的一个样本，

$p(x\,|\,\theta)=\theta x^{\theta-1}$，$0<x<1$．

若取 $\theta$ 的先验分布为伽玛分布，即 $\theta\sim Ga(\alpha, \lambda)$，求 $\theta$ 的后验期望估计．

解：参数 $\theta$ 的先验分布为 $Ga(\alpha, \lambda)$，密度函数为 $\pi(\theta)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\lambda\theta}\mathrm{I}_{\theta>0}$，

因样本 $X_1, \cdots, X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}\mathrm{I}_{0<x_i<1}=\theta^n(x_1\cdots x_n)^{\theta-1}\mathrm{I}_{0<x_1,\cdots,x_n<1}=\theta^n\mathrm{e}^{(\theta-1)\ln(x_1\cdots x_n)}\mathrm{I}_{0<x_1,\cdots,x_n<1}\text{，}$$

则样本 $X_1, \cdots, X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\theta^{n+\alpha-1}\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\mathrm{I}_{0<x_1,\cdots,x_n<1,\,\theta>0}\text{，}$$

样本 $X_1, \cdots, X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\theta^{n+\alpha-1}\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\mathrm{I}_{0<x_1,\cdots,x_n<1}d\theta$$

$$=\frac{\lambda^\alpha}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\cdot\frac{\Gamma(n+\alpha)}{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}\mathrm{I}_{0<x_1,\cdots,x_n<1}\text{，}$$

即参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{m(x_1,\cdots,x_n)}=\frac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\theta^{n+\alpha-1}\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}\mathrm{I}_{\theta>0}\text{，}$$

后验分布仍为伽玛分布 $Ga(n+\alpha, \lambda-\ln(x_1\cdots x_n))$，

因 $E(\theta\,|\,x_1,\cdots,x_n)=\displaystyle\int_0^1\theta\cdot\pi(\theta\,|\,x_1,\cdots,x_n)d\theta=\frac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\int_0^1\theta^{n+\alpha}\mathrm{e}^{-[\lambda-\ln(x_1\cdots x_n)]\theta}d\theta$

$$=\frac{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}\cdot\frac{\Gamma(n+\alpha+1)}{[\lambda-\ln(x_1\cdots x_n)]^{n+\alpha+1}}=\frac{n+\alpha}{\lambda-\ln(x_1\cdots x_n)}\text{，}$$

故参数 $\theta$ 的后验期望估计 $\hat{\theta}_B=\dfrac{n+\alpha}{\lambda-\ln(X_1\cdots X_n)}$．

8． 设 $X_1, \cdots, X_n$ 是来自均匀分布 $U(0, \theta)$ 的样本，$\theta$ 的先验分布是帕雷托（Pareto）分布，密度函数为 $\pi(\theta)=\dfrac{\beta\theta_0^\beta}{\theta^{\beta+1}}$，$\theta>\theta_0$，其中 $\beta, \theta_0$ 是两个已知的常数．

（1）验证：帕雷托分布是 $\theta$ 的共轭先验分布；

（2）求$\theta$的贝叶斯估计．

解：（1）参数$\theta$的先验分布是帕雷托分布，密度函数为$\pi(\theta)=\dfrac{\beta\theta_0^\beta}{\theta^{\beta+1}}\mathrm{I}_{\theta>\theta_0}$，

因样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^n\frac{1}{\theta}\mathrm{I}_{0<x_i<\theta}=\frac{1}{\theta^n}\mathrm{I}_{0<x_1,\cdots,x_n<\theta},$$

则样本$X_1,\cdots,X_n$和参数$\theta$的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\,\theta>\theta_0}=\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0,\,\theta>\max\{x_1,\cdots,x_n,\,\theta_0\}},$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,\,\theta_0\}}^{+\infty}\frac{\beta\theta_0^\beta}{\theta^{n+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\beta\theta_0^\beta\cdot\frac{1}{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}\mathrm{I}_{x_1,\cdots,x_n>0},$$

即参数$\theta$的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}{\theta^{n+\beta+1}}\mathrm{I}_{\theta>\max\{x_1,\cdots,x_n,\,\theta_0\}},$$

后验分布仍为帕雷托分布，其参数为$n+\beta$和$\max\{x_1,\cdots,x_n,\theta_0\}$，

故帕雷托分布是参数$\theta$的共轭先验分布；

（2）因$E(\theta\,|\,x_1,\cdots,x_n)=\displaystyle\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\theta\cdot\pi(\theta\,|\,x_1,\cdots,x_n)d\theta$

$$=\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\frac{(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}}{\theta^{n+\beta}}d\theta$$

$$=(n+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{n+\beta}\cdot\frac{[\max\{x_1,\cdots,x_n,\theta_0\}]^{-(n+\beta)+1}}{n+\beta-1}=\frac{n+\beta}{n+\beta-1}\max\{x_1,\cdots,x_n,\theta_0\},$$

故$\theta$的贝叶斯估计$\hat\theta_B=\dfrac{n+\beta}{n+\beta-1}\max\{X_1,\cdots,X_n,\theta_0\}$．

9．设指数分布$Exp(\theta)$中未知参数$\theta$的先验分布为伽玛分布$Ga(\alpha,\lambda)$，现从先验信息得知：先验均值为0.0002，先验标准差为0.01，试确定先验分布．

解：因伽玛分布$Ga(\alpha,\lambda)$密度函数为$\pi(\theta)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\mathrm{e}^{-\lambda\theta}\mathrm{I}_{\theta>0}$，

则由$E(\theta)=\dfrac{\alpha}{\lambda}=0.0002$，$\mathrm{Var}(\theta)=\dfrac{\alpha}{\lambda^2}=(0.01)^2=0.0001$，解得$\lambda=2$，$\alpha=0.0004$，

故参数$\theta$的先验分布为伽玛分布$Ga(0.0004,2)$．

10．设$X_1,\cdots,X_n$为来自如下幂级数分布的样本，总体分布密度为

$$p(x_1;c,\theta)=cx_1^{c-1}\theta^{-c}\mathrm{I}_{0\le x_1\le\theta}\quad(c>0,\theta>0),$$

（1）证明：若$c$已知，则$\theta$的共轭先验分布为帕雷托分布；

（2）若$\theta$已知，则$c$的共轭先验分布为伽玛分布．

证：样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^n cx_i^{c-1}\theta^{-c}\mathrm{I}_{0<x_i<\theta}=c^n(x_1\cdots x_n)^{c-1}\theta^{-nc}\mathrm{I}_{0<x_1,\cdots,x_n<\theta},$$

（1）设参数$\theta$的先验分布是帕雷托分布，密度函数为$\pi(\theta)=\dfrac{\beta\theta_0^{\beta}}{\theta^{\beta+1}}\mathrm{I}_{\theta>\theta_0}$，

则样本$X_1,\cdots,X_n$和参数$\theta$的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\beta\theta_0^{\beta}c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{0<x_1,\cdots,x_n,\theta_0<\theta}=\frac{\beta\theta_0^{\beta}c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0,\,\theta>\max\{x_1,\cdots,x_n,\theta_0\}}，$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,\theta_0\}}^{+\infty}\frac{\beta\theta_0^{\beta}c^n(x_1\cdots x_n)^{c-1}}{\theta^{nc+\beta+1}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta$$

$$=\beta\theta_0^{\beta}c^n(x_1\cdots x_n)^{c-1}\frac{[\max\{x_1,\cdots,x_n,\theta_0\}]^{-(nc+\beta)}}{nc+\beta}\mathrm{I}_{x_1,\cdots,x_n>0}，$$

即参数$\theta$的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{(nc+\beta)[\max\{x_1,\cdots,x_n,\theta_0\}]^{nc+\beta}}{\theta^{nc+\beta+1}}\mathrm{I}_{\theta>\max\{x_1,\cdots,x_n,\theta_0\}}，$$

后验分布仍为帕雷托分布，其参数为$nc+\beta$和$\max\{x_1,\cdots,x_n,\theta_0\}$，
故帕雷托分布是参数$\theta$的共轭先验分布；

（2）设参数$c$的先验分布为伽玛分布$Ga(\alpha,\lambda)$，密度函数为$\pi(c)=\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}c^{\alpha-1}\mathrm{e}^{-\lambda c}\mathrm{I}_{c>0}$，

则样本$X_1,\cdots,X_n$和参数$\theta$的联合分布为

$$h(x_1,\cdots,x_n,c)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}c^{n+\alpha-1}(x_1\cdots x_n)^{c-1}\mathrm{e}^{-\lambda c}\theta^{-nc}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\,c>0}$$

$$=\frac{\lambda^{\alpha}}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}c^{n+\alpha-1}\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\,c>0}，$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^{+\infty}\frac{\lambda^{\alpha}}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}c^{n+\alpha-1}\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\mathrm{I}_{0<x_1,\cdots,x_n<\theta}d\theta$$

$$=\frac{\lambda^{\alpha}}{\Gamma(\alpha)\cdot(x_1\cdots x_n)}\cdot\frac{\Gamma(n+\alpha)}{[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]^{n+\alpha}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta}，$$

即参数$\theta$的后验分布为

$$\pi(c\,|\,x_1,\cdots,x_n)=\frac{[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]^{n+\alpha}}{\Gamma(n+\alpha)}c^{n+\alpha-1}\mathrm{e}^{-[\lambda+n\ln\theta-\ln(x_1\cdots x_n)]c}\mathrm{I}_{c>0}，$$

后验分布仍为伽玛分布，其参数为$n+\alpha$和$\lambda+n\ln\theta-\ln(x_1\cdots x_n)$，
故伽玛分布是参数$c$的共轭先验分布．

11．某人每天早上在汽车站等公共汽车的时间（单位：min）服从均匀分布$U(0,\theta)$，其中$\theta$未知，假设$\theta$的先验分布为

$$\pi(\theta)=\begin{cases}\dfrac{192}{\theta^4}, & \theta\ge 4;\\ 0, & \theta<4.\end{cases}$$

假如此人在三个早上等车的时间分别为5, 3, 8分钟，求$\theta$后验分布．

解：参数$\theta$的先验分布为$\pi(\theta)=\dfrac{192}{\theta^4}\mathrm{I}_{\theta>4}$，

因样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\mid\theta)=\prod_{i=1}^{n}\frac{1}{\theta}\mathrm{I}_{0<x_i<\theta}=\frac{1}{\theta^n}\mathrm{I}_{0<x_1,\cdots,x_n<\theta},$$

则样本$X_1,\cdots,X_n$和参数$\theta$的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{192}{\theta^{n+4}}\mathrm{I}_{0<x_1,\cdots,x_n<\theta,\,\theta>4}=\frac{192}{\theta^{n+4}}\mathrm{I}_{x_1,\cdots,x_n>0,\,\theta>\max\{x_1,\cdots,x_n,4\}},$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_{\max\{x_1,\cdots,x_n,4\}}^{+\infty}\frac{192}{\theta^{n+4}}\mathrm{I}_{x_1,\cdots,x_n>0}d\theta=\frac{192}{(n+3)[\max\{x_1,\cdots,x_n,4\}]^{n+3}}\mathrm{I}_{x_1,\cdots,x_n>0},$$

即参数$\theta$的后验分布为

$$\pi(\theta\mid x_1,\cdots,x_n)=\frac{(n+3)[\max\{x_1,\cdots,x_n,4\}]^{n+3}}{\theta^{n+4}}\mathrm{I}_{\theta>\max\{x_1,\cdots,x_n,4\}},$$

后验分布仍为帕雷托分布，其参数为$n+3$和$\max\{x_1,\cdots,x_n,4\}$，

因样本观测值为 5, 3, 8，即$\max\{x_1,\cdots,x_n,4\}=8$，$n=3$，

故参数$\theta$的后验分布为帕雷托分布，其参数为 6 和 8，密度函数为

$$\pi(\theta\mid x_1,x_2,x_3)=\frac{6\times8^6}{\theta^7}\mathrm{I}_{\theta>8}.$$

12．从正态分布$N(\theta,\ 2^2)$中随机抽取容量为 100 的样本，又设$\theta$的先验分布为正态分布，证明：不管先验分布的标准差为多少，后验分布的标准差一定小于 1/5.

解：设$\theta$的先验分布为正态分布$N(\mu,\sigma^2)$，根据书上 P336 例 6.5.3 的结论可知，$\theta$的后验分布为

$$N\left(\frac{2^{-2}n\overline{X}+\mu\sigma^{-2}}{2^{-2}n+\sigma^{-2}},\frac{1}{2^{-2}n+\sigma^{-2}}\right)=N\left(\frac{25\overline{X}+\mu\sigma^{-2}}{25+\sigma^{-2}},\frac{1}{25+\sigma^{-2}}\right),$$

故后验分布的标准差为$\sqrt{\dfrac{1}{25+\sigma^{-2}}}<\dfrac{1}{5}$.

13．设随机变量$X$服从负二项分布，其概率分布为

$$f(x\mid p)=\binom{x-1}{k-1}p^k(1-p)^{x-k},\quad x=k,k+1,\cdots,$$

证明其成功概率$p$共轭先验分布族为贝塔分布族.

证：设参数$p$的先验分布是贝塔分布$Be(a,b)$，密度函数为$\pi(p)=\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}\mathrm{I}_{0<p<1}$，

因样本$X_1,\cdots,X_n$的联合条件分布为

$$p(x_1,\cdots,x_n\mid p)=\prod_{i=1}^{n}\binom{x_i-1}{k-1}p^k(1-p)^{x_i-k}=\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot p^{nk}(1-p)^{\sum_{i=1}^{n}x_i-nk},$$

则样本$X_1,\cdots,X_n$和参数$p$的联合分布为

$$h(x_1,\cdots,x_n,p)=\prod_{i=1}^{n}\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{nk+a-1}(1-p)^{\sum_{i=1}^{n}x_i-nk+b-1}\mathrm{I}_{0<p<1},$$

样本$X_1,\cdots,X_n$的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^1\prod_{i=1}^n\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{nk+a-1}(1-p)^{\sum_{i=1}^n x_i-nk+b-1}dp$$

$$=\prod_{i=1}^n\binom{x_i-1}{k-1}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\cdot\frac{\Gamma(nk+a)\cdot\Gamma\left(\sum_{i=1}^n x_i-nk+b\right)}{\Gamma\left(\sum_{i=1}^n x_i+a+b\right)},$$

即参数 $p$ 的后验分布为

$$\pi(p\,|\,x_1,\cdots,x_n)=\frac{\Gamma\left(\sum_{i=1}^n x_i+a+b\right)}{\Gamma(nk+a)\cdot\Gamma\left(\sum_{i=1}^n x_i-nk+b\right)}p^{nk+a-1}(1-p)^{\sum_{i=1}^n x_i-nk+b-1}\mathrm{I}_{0<p<1},$$

后验分布仍为贝塔分布，其参数为 $nk+a$ 和 $\sum_{i=1}^n x_i-nk+b$，

故贝塔分布是参数 $p$ 的共轭先验分布.

14．从一批产品中抽检 100 个，发现 3 个不合格，假定该产品不合格率 $\theta$ 的先验分布为贝塔分布 $Be(2, 200)$，求 $\theta$ 的后验分布.

解：参数 $\theta$ 的先验分布是贝塔分布 $Be(2, 200)$，密度函数为 $\pi(\theta)=\dfrac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta(1-\theta)^{199}\mathrm{I}_{0<\theta<1}$，

因样本 $X_1,\cdots,X_n$ 的联合条件分布为

$$p(x_1,\cdots,x_n\,|\,\theta)=\prod_{i=1}^n\theta^{x_i}(1-\theta)^{1-x_i}=\theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i},$$

则样本 $X_1,\cdots,X_n$ 和参数 $\theta$ 的联合分布为

$$h(x_1,\cdots,x_n,\theta)=\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta^{1+\sum_{i=1}^n x_i}(1-\theta)^{n+199-\sum_{i=1}^n x_i}\mathrm{I}_{0<\theta<1},$$

样本 $X_1,\cdots,X_n$ 的边际分布为

$$m(x_1,\cdots,x_n)=\int_0^1\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\theta^{1+\sum_{i=1}^n x_i}(1-\theta)^{n+199-\sum_{i=1}^n x_i}d\theta$$

$$=\frac{\Gamma(202)}{\Gamma(2)\Gamma(200)}\cdot\frac{\Gamma\left(2+\sum_{i=1}^n x_i\right)\Gamma\left(n+200-\sum_{i=1}^n x_i\right)}{\Gamma(n+202)},$$

即参数 $\theta$ 的后验分布为

$$\pi(\theta\,|\,x_1,\cdots,x_n)=\frac{\Gamma(n+202)}{\Gamma\left(2+\sum_{i=1}^n x_i\right)\Gamma\left(n+200-\sum_{i=1}^n x_i\right)}\theta^{1+\sum_{i=1}^n x_i}(1-\theta)^{n+199-\sum_{i=1}^n x_i}\mathrm{I}_{0<\theta<1},$$

后验分布仍为贝塔分布，其参数为 $2+\sum_{i=1}^n x_i$ 和 $n+200-\sum_{i=1}^n x_i$，

因 $n=100$，$\sum_{i=1}^{n} x_i = 3$，

故参数$\theta$的后验分布为贝塔分布 $Be(5, 297)$，密度函数为

$$\pi(\theta \mid x_1, \cdots, x_n) = \frac{\Gamma(302)}{\Gamma(5)\Gamma(297)} \theta^4 (1-\theta)^{296} \mathrm{I}_{0<\theta<1} .$$

## 习题 6.6

1．某厂生产的化纤强度服从正态分布，长期以来其标准差稳定在$\sigma = 0.85$，现抽取了一个容量为 $n = 25$ 的样本，测定其强度，算得平均值为 $\bar{x} = 2.25$，试求这批化纤平均强度的置信水平为 0.95 的置信区间.

解：已知$\sigma^2$，估计$\mu$，选取枢轴量$U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，置信区间为$\left[\overline{X} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$，

置信度 $1-\alpha = 0.95$，$u_{1-\alpha/2} = u_{0.975} = 1.96$，$\bar{x} = 2.25$，$\sigma = 0.85$，$n = 25$，

故$\mu$ 的 0.95 置信区间为$\left[\bar{x} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right] = \left[2.25 \pm 1.96 \times \dfrac{0.85}{\sqrt{25}}\right] = [1.9168, 2.5832]$.

2．总体 $X \sim N(\mu, \sigma^2)$，$\sigma^2$ 已知，问样本容量 $n$ 取多大时才能保证$\mu$ 的置信水平为 95%的置信区间的长度不大于 $k$.

解：已知$\sigma^2$，估计$\mu$，选取枢轴量$U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，置信区间为$\left[\overline{X} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$，长度为$2u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}$，

置信度 $1-\alpha = 0.95$，$u_{1-\alpha/2} = u_{0.975} = 1.96$，有置信区间的长度$2u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}} = 2 \times 1.96 \times \dfrac{\sigma}{\sqrt{n}} \le k$，

故$\sqrt{n} \ge 3.92 \times \dfrac{\sigma}{k}$，即$n \ge \dfrac{15.3664\sigma^2}{k^2}$.

3．0.50, 1.25, 0.80, 2.00 是取自总体 $X$ 的样本，已知 $Y = \ln X$ 服从正态分布 $N(\mu, 1)$.

（1）求$\mu$ 的置信水平为 95%的置信区间；

（2）求 $X$ 的数学期望的置信水平为 95%的置信区间.

解：（1）已知$\sigma^2$，估计$\mu$，选取枢轴量$U = \dfrac{\overline{Y}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，置信区间为$\left[\overline{Y} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$，

置信度 $1-\alpha = 0.95$，$u_{1-\alpha/2} = u_{0.975} = 1.96$，$\sigma = 1$，$n = 4$，

$$\bar{y} = \frac{1}{4}(\ln 0.50 + \ln 1.25 + \ln 0.80 + \ln 2.00) = 0,$$

故$\mu$ 的 95%置信区间为$\left[\bar{y} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right] = \left[0 \pm 1.96 \times \dfrac{1}{\sqrt{4}}\right] = [-0.98, 0.98]$；

（2）因 $Y = \ln X$ 服从正态分布 $N(\mu, 1)$，有 $X = \mathrm{e}^Y$，且 $Y$ 的密度函数为 $p(y) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu)^2}{2}}$，

则$E(X) = \int_{-\infty}^{+\infty} \mathrm{e}^y \cdot \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(y-\mu)^2}{2}} dy = \dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{y^2-2\mu y+\mu^2-2y}{2}} dy$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{y^2-2(\mu+1)y+(\mu+1)^2-2\mu-1}{2}}dy=e^{\mu+\frac{1}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-\mu-1)^2}{2}}dy=e^{\mu+\frac{1}{2}},$$

故 $E(X)$ 的 95%置信区间为 $[e^{-0.98+0.5}, e^{0.98+0.5}]=[0.6188, 4.3929]$.

4. 用一个仪表测量某一物理量 9 次，得样本均值 $\bar{x}=56.32$，样本标准差 $s=0.22$.

（1）测量标准差$\sigma$大小反映了测量仪表的精度，试求$\sigma$的置信水平为 0.95 置信区间；

（2）求该物理量真值的置信水平为 0.99 的置信区间.

解：（1）估计 $\sigma^2$，选取枢轴量 $\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，置信区间为 $\left[\frac{(n-1)\cdot S^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)\cdot S^2}{\chi^2_{\alpha/2}(n-1)}\right]$,

置信度 $1-\alpha=0.95$，$n=9$，$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(8)=2.1797$，$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.975}(8)=17.5345$，

$s=0.22$，

故 $\sigma^2$ 的 0.95 置信区间为 $\left[\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2}(n-1)}\right]=\left[\frac{8\times0.22^2}{17.5345}, \frac{8\times0.22^2}{2.1797}\right]=[0.0221, 0.1776]$，

即 $\sigma$ 的 0.95 置信区间为 $[\sqrt{0.0221}, \sqrt{0.1776}]=[0.1486, 0.4215]$.

（2）未知 $\sigma^2$，估计$\mu$，选取枢轴量 $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$，置信区间为 $\left[\bar{X}\pm t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right]$,

置信度 $1-\alpha=0.99$，$n=9$，$t_{1-\alpha/2}(n-1)=t_{0.995}(8)=3.3554$，$\bar{x}=56.32$，$s=0.22$，

故$\mu$ 的 0.99 置信区间为 $\left[\bar{x}\pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right]=\left[56.32\pm3.3554\times\frac{0.22}{\sqrt{9}}\right]=[56.0739, 56.5661]$.

5. 已知某种材料的抗压强度 $X\sim N(\mu, \sigma^2)$，现随机地抽取 10 个试件进行抗压试验，测得数据如下：

482 493 457 471 510 446 435 418 394 469

（1）求平均抗压强度$\mu$ 的置信水平为 95%的置信区间；

（2）若已知$\sigma=30$，求平均抗压强度$\mu$ 的置信水平为 95%的置信区间；

（3）求$\sigma$ 的置信水平为 95%的置信区间.

解：（1）未知 $\sigma^2$，估计$\mu$，选取枢轴量 $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$，置信区间为 $\left[\bar{X}\pm t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right]$,

置信度 $1-\alpha=0.95$，$n=10$，$t_{1-\alpha/2}(n-1)=t_{0.975}(9)=2.2622$，$\bar{x}=457.5$，$s=35.2176$，

故$\mu$ 的 95%置信区间 $\left[\bar{x}\pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right]=\left[457.5\pm2.2622\times\frac{35.2176}{\sqrt{10}}\right]=[432.3064, 482.6936]$；

（2）已知$\sigma^2$，估计$\mu$，选取枢轴量 $U=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，置信区间为 $\left[\bar{X}\pm u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$,

置信度 $1-\alpha=0.95$，$u_{1-\alpha/2}=u_{0.975}=1.96$，$\bar{x}=457.5$，$\sigma=30$，$n=10$，

故$\mu$ 的 95%置信区间为 $\left[\bar{x}\pm u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]=\left[457.5\pm1.96\times\frac{30}{\sqrt{10}}\right]=[438.9058, 476.0942]$；

（3）估计 $\sigma^2$，选取枢轴量 $\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，置信区间为 $\left[\frac{(n-1)\cdot S^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)\cdot S^2}{\chi^2_{\alpha/2}(n-1)}\right]$,

置信度 $1-\alpha=0.95$，$n=10$，$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(9)=2.7004$，$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.975}(9)=19.0228$，

$s=35.2176$，

故 $\sigma^2$ 的 0.95 置信区间为

$$\left[\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2}(n-1)},\ \frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2}(n-1)}\right]=\left[\frac{9\times 35.2176^2}{19.0228},\ \frac{9\times 35.2176^2}{2.7004}\right]=[586.7958,\ 4133.6469]，$$

即 $\sigma$ 的 0.95 置信区间为 $[\sqrt{586.7958},\ \sqrt{4133.6469}]=[24.2239, 64.2934]$.

6. 在一批货物中随机抽取 80 件，发现有 11 件不合格品，试求这批货物的不合格品率的置信水平为 0.90 的置信区间.

解：大样本，估计概率 $p$，选取枢轴量 $U=\dfrac{\overline{X}-p}{\sqrt{\dfrac{p(1-p)}{n}}}\dot{\sim} N(0,1)$，

置信区间为 $\dfrac{1}{1+u^2_{1-\alpha/2}/n}\left[\overline{X}+\dfrac{u^2_{1-\alpha/2}}{2n}\pm u_{1-\alpha/2}\sqrt{\dfrac{\overline{X}(1-\overline{X})}{n}+\dfrac{u^2_{1-\alpha/2}}{4n^2}}\right]$，

置信度 $1-\alpha=0.90$，$u_{1-\alpha/2}=u_{0.95}=1.645$，$n=80$，$\bar{x}=\dfrac{11}{80}=0.1375$，

故 $p$ 的 0.90 置信区间

$$\frac{1}{1+u^2_{1-\alpha/2}/n}\left[\bar{x}+\frac{u^2_{1-\alpha/2}}{2n}\pm u_{1-\alpha/2}\sqrt{\frac{\bar{x}(1-\bar{x})}{n}+\frac{u^2_{1-\alpha/2}}{4n^2}}\right]$$

$$=\frac{1}{1+1.645^2/80}\left[0.1375+\frac{1.645^2}{160}\pm 1.645\times\sqrt{\frac{0.1375\times 0.8625}{80}+\frac{1.645^2}{4\times 80^2}}\right]=[0.0859, 0.2128].$$

注：$p$ 的 0.90 近似置信区间

$$\left[\bar{x}\pm u_{1-\alpha/2}\sqrt{\frac{\bar{x}(1-\bar{x})}{n}}\right]=\left[0.1375\pm 1.645\times\sqrt{\frac{0.1375\times 0.8625}{80}}\right]=[0.0742, 0.2008]；$$

$p$ 的 0.90 修正置信区间（修正频率 $\bar{x}^*=\dfrac{11+2}{80+4}=0.1548$）

$$\left[\bar{x}^*\pm u_{1-\alpha/2}\sqrt{\frac{\bar{x}^*(1-\bar{x}^*)}{n+4}}\right]=\left[0.1548\pm 1.645\times\sqrt{\frac{0.1548\times 0.8452}{84}}\right]=[0.0898, 0.2197].$$

7. 设 $X_1,\cdots,X_n$ 是来自泊松分布 $P(\lambda)$ 的样本，证明：$\lambda$ 的近似 $1-\alpha$ 置信区间为

$$\left[\frac{2\overline{X}+\frac{1}{n}u^2_{1-\alpha/2}-\sqrt{\left(2\overline{X}+\frac{1}{n}u^2_{1-\alpha/2}\right)^2-4\overline{X}^2}}{2},\frac{2\overline{X}+\frac{1}{n}u^2_{1-\alpha/2}+\sqrt{\left(2\overline{X}+\frac{1}{n}u^2_{1-\alpha/2}\right)^2-4\overline{X}^2}}{2}\right].$$

证：总体 $X\sim P(\lambda)$，有 $n\overline{X}=X_1+\cdots+X_n\sim P(n\lambda)$，$E(\overline{X})=\lambda$，$\mathrm{Var}(\overline{X})=\dfrac{\lambda}{n}$，当 $n$ 很大时，$\overline{X}\dot{\sim} N\left(\lambda,\dfrac{\lambda}{n}\right)$，

选取枢轴量$U=\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\sim N(0,1)$，置信度为 $1-\alpha$，即$P\left\{-u_{1-\alpha/2}\le\dfrac{\overline{X}-\lambda}{\sqrt{\lambda/n}}\le u_{1-\alpha/2}\right\}=1-\alpha$，

则$-u_{1-\alpha/2}\sqrt{\dfrac{\lambda}{n}}\le\overline{X}-\lambda\le u_{1-\alpha/2}\sqrt{\dfrac{\lambda}{n}}$，即$(\overline{X}-\lambda)^2\le u_{1-\alpha/2}^2\cdot\dfrac{\lambda}{n}$，$\lambda^2-\left(2\overline{X}+\dfrac{1}{n}u_{1-\alpha/2}^2\right)\lambda+\overline{X}^2\le 0$，

解得$\dfrac{2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2-\sqrt{\left(2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\right)^2-4\overline{X}^2}}{2}\le\lambda\le\dfrac{2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2+\sqrt{\left(2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\right)^2-4\overline{X}^2}}{2}$，

置信区间为$\left[\dfrac{2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2-\sqrt{\left(2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\right)^2-4\overline{X}^2}}{2},\dfrac{2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2+\sqrt{\left(2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\right)^2-4\overline{X}^2}}{2}\right]$.

8． 某商店某种商品的月销售量服从泊松分布，为合理进货，必须了解销售情况．现记录了该商店过去的一些销售量，数据如下：

| 月销售量 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|
| 月份数 | 1 | 6 | 13 | 12 | 9 | 4 | 2 | 1 |

试求平均月销售量的置信水平为 0.95 的置信区间．

解：估计泊松分布的参数$\lambda$，由第 7 题的结论可知$\lambda$ 的近似 $1-\alpha$ 置信区间为

$$\left[\frac{2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\pm\sqrt{\left(2\overline{X}+\frac{1}{n}u_{1-\alpha/2}^2\right)^2-4\overline{X}^2}}{2}\right]=\left[\overline{X}+\frac{1}{2n}u_{1-\alpha/2}^2\pm\sqrt{\left(\overline{X}+\frac{1}{2n}u_{1-\alpha/2}^2\right)^2-\overline{X}^2}\right],$$

置信度 $1-\alpha=0.95$，$u_{1-\alpha/2}=u_{0.975}=1.96$，$\bar{x}=11.9792$，$n=48$，

故$\lambda$ 的 0.95 置信区间

$$\left[\bar{x}+\frac{1}{2n}u_{1-\alpha/2}^2\pm\sqrt{\left(\bar{x}+\frac{1}{2n}u_{1-\alpha/2}^2\right)^2-\bar{x}^2}\right]$$

$$=\left[11.9792+\frac{1.96^2}{2\times 48}\pm\sqrt{\left(11.9792+\frac{1.96^2}{2\times 48}\right)^2-11.9792^2}\right]=[11.0392,12.9992].$$

9． 设从总体$X\sim N(\mu_1,\sigma_1^2)$和总体$Y\sim N(\mu_2,\sigma_2^2)$中分别抽取容量为 $n_1=10, n_2=15$ 的独立样本，可计算得$\bar{x}=82,\ s_x^2=56.5,\ \bar{y}=76,\ s_y^2=52.4$.

（1）若已知$\sigma_1^2=64,\ \sigma_2^2=49$，求$\mu_1-\mu_2$的置信水平为 95%的置信区间；

（2）若已知$\sigma_1^2=\sigma_2^2$，求$\mu_1-\mu_2$的置信水平为 95%的置信区间；

（3）若对$\sigma_1^2,\sigma_2^2$一无所知，求$\mu_1-\mu_2$的置信水平为 95%的近似置信区间；

（4）求$\sigma_1^2/\sigma_2^2$ 的置信水平为 95%的置信区间．

解：（1）已知$\sigma_1^2,\sigma_2^2$，估计$\mu_1-\mu_2$，选取枢轴量$U=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$，

置信区间为$\left[\overline{X}-\overline{Y}\pm u_{1-\alpha/2}\cdot\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}\right]$，

置信度 $1-\alpha=0.95$，$u_{1-\alpha/2}=u_{0.975}=1.96$，$\bar{x}=82,\ \bar{y}=76,\ \sigma_1^2=64,\ \sigma_2^2=49$，$n_1=10$，$n_2=15$，

故$\mu_1-\mu_2$的 95%置信区间为

$$\left[\bar{x}-\bar{y}\pm u_{1-\alpha/2}\cdot\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]=\left[82-76\pm1.96\times\sqrt{\frac{64}{10}+\frac{49}{15}}\right]=[-0.0939,12.0939]；$$

（2）未知$\sigma_1^2,\sigma_2^2$，但$\sigma_1^2=\sigma_2^2$，估计$\mu_1-\mu_2$，选取枢轴量$T=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$，

置信区间为$\left[\overline{X}-\overline{Y}\pm t_{1-\alpha/2}(n_1+n_2-2)\cdot S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\right]$，

置信度 $1-\alpha=0.95$，$n_1=10$，$n_2=15$，$t_{1-\alpha/2}(n_1+n_2-2)=t_{0.975}(23)=2.0687$，

$\bar{x}=82,\ s_x^2=56.5,\ \bar{y}=76,\ s_y^2=52.4$，有$s_w=\sqrt{\dfrac{9\times56.5+14\times52.4}{23}}=7.3488$，

故$\mu_1-\mu_2$的 95%置信区间为

$$\left[\bar{x}-\bar{y}\pm t_{1-\alpha/2}(n_1+n_2-2)\cdot s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=\left[82-76\pm2.0687\times7.3488\times\sqrt{\frac{1}{10}+\frac{1}{15}}\right]$$

$$=[-0.2063,\ 12.2063];$$

（3）未知$\sigma_1^2,\sigma_2^2$，估计$\mu_1-\mu_2$，

选取枢轴量$T=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}}}\ \dot{\sim}\ t(l_0)$，$l_0$是最接近$l=\dfrac{\left(\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}\right)^2}{\dfrac{S_x^4}{n_1^2(n_1-1)}+\dfrac{S_y^4}{n_2^2(n_2-1)}}$的整数，

近似置信区间为$\mu_1-\mu_2\in\left[\overline{X}-\overline{Y}\pm t_{1-\alpha/2}(l_0)\cdot\sqrt{\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}}\right]$，

因$n_1=10$，$n_2=15$，$s_x^2=56.5,\ s_y^2=52.4$，有$l=\dfrac{\left(\dfrac{56.5}{10}+\dfrac{52.4}{15}\right)^2}{\dfrac{56.5^2}{10^2\times9}+\dfrac{52.4^2}{15^2\times14}}=18.9201$，即取$l_0=19$，

置信度为 $1-\alpha=0.95$，$t_{1-\alpha/2}(l_0)=t_{0.975}(19)=2.0930$，$\bar{x}=82,\ s_x^2=56.5,\ \bar{y}=76,\ s_y^2=52.4$，

故$\mu_1-\mu_2$的 95%置信区间为

$$\left[\bar{x}-\bar{y}\pm t_{1-\alpha/2}(l_0)\cdot\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}\right]=\left[82-76\pm 2.0930\times\sqrt{\frac{56.5}{10}+\frac{52.4}{15}}\right]=[-0.3288, 12.3288]；$$

（4）估计方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$，选取枢轴量 $F=\dfrac{S_x^2/\sigma_1^2}{S_y^2/\sigma_2^2}\sim F(n_1-1,\ n_2-1)$，

置信区间为 $\left[\dfrac{S_x^2}{S_y^2}\cdot\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)},\ \dfrac{S_x^2}{S_y^2}\cdot\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)}\right]$，

置信度 $1-\alpha=0.95$，$n_1=10$，$n_2=15$，$F_{1-\alpha/2}(n_1-1,n_2-1)=F_{0.975}(9,14)=3.21$，

$$F_{\alpha/2}(n_1-1,n_2-1)=F_{0.025}(9,14)=\frac{1}{F_{0.975}(14,9)}=\frac{1}{3.80}，\quad s_x^2=56.5,\ s_y^2=52.4，$$

故 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的 95%置信区间为

$$\left[\frac{s_x^2}{s_y^2}\cdot\frac{1}{F_{0.975}(9,14)},\ \frac{s_x^2}{s_y^2}\cdot\frac{1}{F_{0.025}(9,14)}\right]=\left[\frac{56.50}{52.4}\times\frac{1}{3.21},\ \frac{56.50}{52.4}\times 3.80\right]=[0.3359, 4.0973].$$

10．假设人体身高服从正态分布，今抽测甲、乙两地区 18 岁~25 岁女青年身高得数据如下：甲地区抽取 10 名，样本均值 1.64 m，样本标准差 0.2 m；乙地区抽取 10 名，样本均值 1.62 m，样本标准差 0.4 m.

（1）两正态总体方差比的置信水平为 95%的置信区间；

（2）两正态总体均值差的置信水平为 95%的置信区间.

解：（1）估计方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$，选取枢轴量 $F=\dfrac{S_x^2/\sigma_1^2}{S_y^2/\sigma_2^2}\sim F(n_1-1,\ n_2-1)$，

置信区间为 $\left[\dfrac{S_x^2}{S_y^2}\cdot\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)},\ \dfrac{S_x^2}{S_y^2}\cdot\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)}\right]$，

置信度 $1-\alpha=0.95$，$n_1=10$，$n_2=10$，$F_{1-\alpha/2}(n_1-1,n_2-1)=F_{0.975}(9,9)=4.03$，

$s_x=0.2$，$s_y=0.4$，

故 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的 95%置信区间为

$$\left[\frac{s_x^2}{s_y^2}\cdot\frac{1}{F_{0.975}(9,9)},\ \frac{s_x^2}{s_y^2}\cdot\frac{1}{F_{0.025}(9,9)}\right]=\left[\frac{0.2^2}{0.4^2}\times\frac{1}{4.03},\ \frac{0.2^2}{0.4^2}\times 4.03\right]=[0.0620, 1.0075]；$$

（2）未知 $\sigma_1^2,\sigma_2^2$，估计 $\mu_1-\mu_2$，

选取枢轴量 $T=\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}}}\dot{\sim} t(l_0)$，$l_0$ 是最接近 $l=\dfrac{\left(\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}\right)^2}{\dfrac{S_x^4}{n_1^2(n_1-1)}+\dfrac{S_y^4}{n_2^2(n_2-1)}}$ 的整数，

近似置信区间为 $\mu_1-\mu_2\in\left[\overline{X}-\overline{Y}\pm t_{1-\alpha/2}(l_0)\cdot\sqrt{\dfrac{S_x^2}{n_1}+\dfrac{S_y^2}{n_2}}\right]$，

因 $n_1=10$，$n_2=10$，$s_x=0.2$，$s_y=0.4$，有 $l=\dfrac{\left(\dfrac{0.2^2}{10}+\dfrac{0.4^2}{10}\right)^2}{\dfrac{0.2^4}{10^2\times 9}+\dfrac{0.4^4}{10^2\times 9}}=13.2353$，即取 $l_0=13$，

置信度为 $1-\alpha=0.95$，$t_{1-\alpha/2}(l_0)=t_{0.975}(13)=2.1604$，$\bar{x}=1.64,\ s_x=0.2,\ \bar{y}=1.62,\ s_y=0.4$，

故 $\mu_1-\mu_2$ 的 95%置信区间为

$$\left[\bar{x}-\bar{y}\pm t_{1-\alpha/2}(l_0)\cdot\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}\right]=\left[1.64-1.62\pm 2.1604\times\sqrt{\frac{0.2^2}{10}+\frac{0.4^2}{10}}\right]=[-0.2855,0.3255].$$

11．设总体 $X$ 的密度函数为 $\lambda\mathrm{e}^{-\lambda x}\mathrm{I}_{x>0}$，其中 $\lambda>0$ 为未知参数，$X_1,\cdots,X_n$ 为抽自此总体的简单随机样本，求 $\lambda$ 的置信水平为 $1-\alpha$ 的置信区间.

解：总体 $X$ 服从指数分布 $Exp(\lambda)$，有 $Y=2\lambda X\sim Exp\left(\dfrac{1}{2}\right)=Ga\left(1,\dfrac{1}{2}\right)=\chi^2(2)$，$n\overline{Y}=Y_1+\cdots+Y_n\sim\chi^2(2n)$，

选取枢轴量 $\chi^2=2n\lambda\overline{X}\sim\chi^2(2n)$，置信度为 $1-\alpha$，即 $P\{\chi^2_{\alpha/2}(2n)\le 2n\lambda\overline{X}\le\chi^2_{1-\alpha/2}(2n)\}=1-\alpha$，

则 $\chi^2_{\alpha/2}(2n)\le 2n\lambda\overline{X}\le\chi^2_{1-\alpha/2}(2n)$，即 $\dfrac{\chi^2_{\alpha/2}(2n)}{2n\overline{X}}\le\lambda\le\dfrac{\chi^2_{1-\alpha/2}(2n)}{2n\overline{X}}$，

故 $\lambda$ 的置信水平为 $1-\alpha$ 的置信区间为 $\left[\dfrac{\chi^2_{\alpha/2}(2n)}{2n\overline{X}},\dfrac{\chi^2_{1-\alpha/2}(2n)}{2n\overline{X}}\right]$.

12．设某电子产品的寿命服从指数分布，其密度函数为 $\lambda\mathrm{e}^{-\lambda x}\mathrm{I}_{x>0}$，现从此批产品中抽取容量为 9 的样本，测得寿命为（单位：千小时）

15，45，50，53，60，65，70，83，90，

求平均寿命 $1/\lambda$ 的置信水平为 0.9 的置信区间和置信上、下限.

解：估计指数分布的参数 $\lambda$，由第 11 题的结论可知 $\lambda$ 的 $1-\alpha$ 置信区间为 $\left[\dfrac{\chi^2_{\alpha/2}(2n)}{2n\overline{X}},\dfrac{\chi^2_{1-\alpha/2}(2n)}{2n\overline{X}}\right]$，

则平均寿命 $1/\lambda$ 的 $1-\alpha$ 置信区间为 $\left[\dfrac{2n\overline{X}}{\chi^2_{1-\alpha/2}(2n)},\dfrac{2n\overline{X}}{\chi^2_{\alpha/2}(2n)}\right]$，

单侧置信上、下限分别为 $\dfrac{2n\overline{X}}{\chi^2_{\alpha}(2n)}$、$\dfrac{2n\overline{X}}{\chi^2_{1-\alpha}(2n)}$，

置信度 $1-\alpha=0.9$，$n=9$，$\chi^2_{\alpha/2}(2n)=\chi^2_{0.05}(18)=9.3905$，$\chi^2_{1-\alpha/2}(2n)=\chi^2_{0.95}(18)=28.8693$，$\bar{x}=59$，

$\chi^2_{\alpha}(2n)=\chi^2_{0.1}(18)=10.8649$，$\chi^2_{1-\alpha}(2n)=\chi^2_{0.9}(18)=25.9894$，

故平均寿命 $1/\lambda$ 的 0.9 置信区间为

$$\left[\frac{2n\overline{X}}{\chi^2_{1-\alpha/2}(2n)}, \frac{2n\overline{X}}{\chi^2_{\alpha/2}(2n)}\right] = \left[\frac{2\times 9\times 59}{28.8693}, \frac{2\times 9\times 59}{9.3905}\right] = [36.7865, 113.0930]\text{；}$$

单侧置信上、下限分别为

$$\frac{2n\overline{X}}{\chi^2_{\alpha}(2n)} = \frac{2\times 9\times 59}{10.8649} = 97.7460\text{，}\quad \frac{2n\overline{X}}{\chi^2_{1-\alpha}(2n)} = \frac{2\times 9\times 59}{10.8649} = 40.8628\text{.}$$

13．设总体 $X$ 的密度函数为

$$p(x;\theta) = \frac{1}{\pi[1+(x-\theta)^2]},\quad -\infty < x < +\infty,\quad -\infty < \theta < +\infty\text{，}$$

$X_1, \cdots, X_n$ 为抽自此总体的简单随机样本，求位置参数 $\theta$ 的置信水平近似为 $1-\alpha$ 的置信区间.

解：总体 $X$ 服从柯西分布，根据书上 P276 例 5.3.10 的结论可知，样本中位数 $m_{0.5} \sim N\left(\theta, \dfrac{\pi^2}{4n}\right)$，

选取枢轴量 $U = \dfrac{m_{0.5}-\theta}{\pi/(2\sqrt{n})} \,\dot\sim\, N(0,1)$，置信度为 $1-\alpha$，即 $P\left\{-u_{1-\alpha/2} \le \dfrac{m_{0.5}-\theta}{\pi/(2\sqrt{n})} \le u_{1-\alpha/2}\right\} = 1-\alpha$，

则 $-u_{1-\alpha/2} \le \dfrac{m_{0.5}-\theta}{\pi/(2\sqrt{n})} \le u_{1-\alpha/2}$，即 $m_{0.5} - u_{1-\alpha/2}\dfrac{\pi}{2\sqrt{n}} \le \theta \le m_{0.5} + u_{1-\alpha/2}\dfrac{\pi}{2\sqrt{n}}$，

故 $\theta$ 的置信水平为 $1-\alpha$ 的近似置信区间为 $\left[m_{0.5} - u_{1-\alpha/2}\dfrac{\pi}{2\sqrt{n}}, m_{0.5} + u_{1-\alpha/2}\dfrac{\pi}{2\sqrt{n}}\right]$.

注：因柯西分布数学期望不存在，由样本均值构造枢轴量得到的置信区间不是一个好的估计，
总体 $X$ 服从柯西分布 $Ch(1, \theta)$，根据书上习题 4.2 第 11 题的结论可知，柯西分布具有可加性，

则 $n\overline{X} = X_1 + \cdots + X_n \sim Ch(n, n\theta)$，有 $Y = n\overline{X} - n\theta \sim Ch(n, 0)$，其密度函数与分布函数分别为

$$p_Y(y) = \frac{n}{\pi(n^2+y^2)}\text{，}\quad F_Y(y) = \int_{-\infty}^{y}\frac{n}{\pi(n^2+t^2)} = \frac{1}{\pi}\arctan\frac{t}{n}\bigg|_{-\infty}^{y} = \frac{1}{2} + \frac{1}{\pi}\arctan\frac{y}{n}\text{，}$$

可得其 $p$ 分位数 $y_p$ 满足 $F_Y(y_p) = \dfrac{1}{2} + \dfrac{1}{\pi}\arctan\dfrac{y_p}{n} = p$，即 $y_p = n\tan\left(\pi p - \dfrac{\pi}{2}\right)$，

选取枢轴量 $Y = n\overline{X} - n\theta \sim Ch(n, 0)$，置信度为 $1-\alpha$，即 $P\{y_{\alpha/2} \le n\overline{X} - n\theta \le y_{1-\alpha/2}\} = 1-\alpha$，

则 $y_{\alpha/2} = -n\tan\dfrac{\pi(1-\alpha)}{2} \le n\overline{X} - n\theta \le y_{1-\alpha/2} = n\tan\dfrac{\pi(1-\alpha)}{2}$，即 $\overline{X} - \tan\dfrac{\pi(1-\alpha)}{2} \le \theta \le \overline{X} + \tan\dfrac{\pi(1-\alpha)}{2}$，

故 $\theta$ 的置信水平为 $1-\alpha$ 的置信区间为 $\left[\overline{X} - \tan\dfrac{\pi(1-\alpha)}{2}, \overline{X} + \tan\dfrac{\pi(1-\alpha)}{2}\right]$.

但是该置信区间长度 $2\tan\dfrac{\pi(1-\alpha)}{2}$ 与样本容量 $n$ 无关，不会随 $n$ 的增加而缩短，不是一个好的估计.

14．设 $X_1, \cdots, X_n$ 为抽自正态总体 $N(\mu, 16)$ 的简单随机样本，为使得 $\mu$ 的置信水平为 $1-\alpha$ 的置信区间的长度不大于给定的 $L$，试问样本容量 $n$ 至少要多少？

解：已知 $\sigma^2$，估计 $\mu$，选取枢轴量 $U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，置信区间为 $\left[\overline{X} \pm u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$，长度为 $2u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}$，

因$\sigma^2 = 16$，有 $2u_{1-\alpha/2}\dfrac{4}{\sqrt{n}} \le L$，

故 $\sqrt{n} \ge \dfrac{8u_{1-\alpha/2}}{L}$，即 $n \ge \dfrac{64u_{1-\alpha/2}^2}{L^2}$．

15．设 $X_1, \cdots, X_n$ 为抽自正态总体 $N(\mu, \sigma^2)$的简单随机样本．试证

$$[\overline{X} - (\mu + k\sigma)] \Big/ \left[\sum_{i=1}^{n}(X_i - \overline{X})^2\right]^{1/2}$$

为枢轴量，其中 $k$ 为已知常数．

证：因 $\dfrac{\overline{X} - (\mu + k\sigma)}{\left[\sum\limits_{i=1}^{n}(X_i - \overline{X})^2\right]^{1/2}} = \dfrac{\overline{X} - (\mu + k\sigma)}{[(n-1)S^2]^{1/2}} = \dfrac{\overline{X} - \mu}{S\sqrt{n-1}} - \dfrac{k\sigma}{[(n-1)S^2]^{1/2}} = \sqrt{n(n-1)}\dfrac{\overline{X} - \mu}{S/\sqrt{n}} - \dfrac{k}{\left[\dfrac{(n-1)S^2}{\sigma^2}\right]^{1/2}}$，

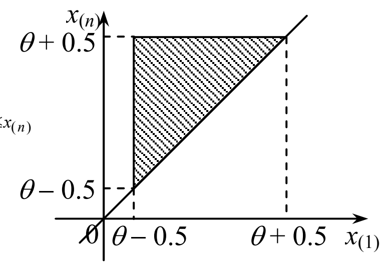
且 $\dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$，$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，分布都与未知参数$\mu, \sigma^2$无关，

故$[\overline{X} - (\mu + k\sigma)] \Big/ \left[\sum\limits_{i=1}^{n}(X_i - \overline{X})^2\right]^{1/2}$ 的分布与未知参数$\mu, \sigma^2$无关，即为枢轴量．

16．设 $X_1, \cdots, X_n$ 是来自 $U(\theta - 1/2, \theta + 1/2)$的样本，求$\theta$的置信水平为 $1 - \alpha$ 的置信区间（提示：证明 $\dfrac{X_{(n)} + X_{(1)}}{2} - \theta$ 为枢轴量，并求出对应的密度函数）．

证：因总体 $X$ 的密度函数与分布函数分别为

$$p(x) = \mathrm{I}_{\theta - 0.5 < x < \theta + 0.5}，\quad F(x) = \begin{cases} 0, & x < \theta - 0.5; \\ x - \theta + 0.5, & \theta - 0.5 \le x < \theta + 0.5; \\ 1, & x \ge \theta + 0.5. \end{cases}$$




则$(X_{(1)}, X_{(n)})$的联合密度函数为

$$p_{1n}(x_{(1)}, x_{(n)}) = n(n-1)[F(x_{(n)}) - F(x_{(1)})]^{n-2} p(x_{(1)}) p(x_{(n)}) \cdot \mathrm{I}_{x_{(1)} \le x_{(n)}}$$

$$= n(n-1)(x_{(n)} - x_{(1)})^{n-2} \mathrm{I}_{\theta - 0.5 < x_{(1)} \le x_{(n)} < \theta + 0.5}，$$

由卷积公式得 $U = X_{(1)} + X_{(n)}$的密度函数，

当 $2\theta - 1 < u < 2\theta$ 时，

$$p_U(u) = \int_{\theta - \frac{1}{2}}^{\frac{u}{2}} n(n-1)[(u - x_{(1)}) - x_{(1)}]^{n-2} dx_{(1)} = -\frac{n}{2}(u - 2x_{(1)})^{n-1}\Big|_{\theta - \frac{1}{2}}^{\frac{u}{2}} = \frac{n}{2}(u - 2\theta + 1)^{n-1}，$$

当 $2\theta \le u < 2\theta + 1$ 时，

$$p_U(u) = \int_{u - \theta - \frac{1}{2}}^{\frac{u}{2}} n(n-1)[(u - x_{(1)}) - x_{(1)}]^{n-2} dx_{(1)} = -\frac{n}{2}(u - 2x_{(1)})^{n-1}\Big|_{u - \theta - \frac{1}{2}}^{\frac{u}{2}} = \frac{n}{2}(2\theta + 1 - u)^{n-1}，$$

当 $u \le 2\theta - 1$ 或 $u \ge 2\theta + 1$ 时，$p_U(u) = 0$，

令 $Y = \dfrac{U}{2} - \theta = \dfrac{X_{(n)} + X_{(1)}}{2} - \theta$，$Y$ 的密度函数与分布函数分别为

$$p_Y(y)=2p_U(2y+2\theta)=\begin{cases}n(1+2y)^{n-1}, & -0.5<y<0;\\ n(1-2y)^{n-1}, & 0\le y<0.5;\\ 0, & 其他.\end{cases}\quad F_Y(y)=\begin{cases}0, & y<-0.5;\\ \frac{1}{2}(1+2y)^n, & -0.5\le y<0;\\ 1-\frac{1}{2}(1-2y)^n, & 0\le y<0.5;\\ 1, & y\ge 0.5.\end{cases}$$

分布与未知参数$\theta$无关，$Y$为枢轴量，

当$p<0.5$时，其$p$分位数$y_p$满足$F_Y(y_p)=\frac{1}{2}(1+2y_p)^n=p$，即$y_p=\frac{(2p)^{\frac{1}{n}}-1}{2}$，

当$p\ge 0.5$时，其$p$分位数$y_p$满足$F_Y(y_p)=1-\frac{1}{2}(1-2y_p)^n=p$，即$y_p=\frac{1-[2(1-p)]^{\frac{1}{n}}}{2}$，

选取枢轴量$Y=\frac{X_{(n)}+X_{(1)}}{2}-\theta$，置信度为 $1-\alpha$，即$P\left\{y_{\alpha/2}\le\frac{X_{(n)}+X_{(1)}}{2}-\theta\le y_{1-\alpha/2}\right\}=1-\alpha$，

则$y_{\alpha/2}=\frac{\alpha^{\frac{1}{n}}-1}{2}\le\frac{X_{(n)}+X_{(1)}}{2}-\theta\le y_{1-\alpha/2}=\frac{1-\alpha^{\frac{1}{n}}}{2}$，即$\frac{X_{(n)}+X_{(1)}}{2}-\frac{1-\alpha^{\frac{1}{n}}}{2}\le\theta\le\frac{X_{(n)}+X_{(1)}}{2}+\frac{1-\alpha^{\frac{1}{n}}}{2}$，

故$\theta$的置信水平为 $1-\alpha$ 的置信区间为$\left[\frac{X_{(n)}+X_{(1)}}{2}-\frac{1-\alpha^{\frac{1}{n}}}{2},\frac{X_{(n)}+X_{(1)}}{2}+\frac{1-\alpha^{\frac{1}{n}}}{2}\right]$.

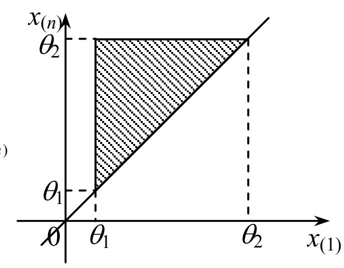
17．设$X_1,\cdots,X_n$为抽自均匀分布$U(\theta_1,\theta_2)$的简单随机样本，记$X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$为其次序统计量．求：

（1）$\theta_2-\theta_1$的置信水平为 $1-\alpha$ 的置信区间；

（2）求$\frac{\theta_2+\theta_1}{2}$的置信水平为 $1-\alpha$ 的置信区间．

解：因总体$X$的密度函数与分布函数分别为

$$p(x)=\frac{1}{\theta_2-\theta_1}\mathrm{I}_{\theta_1<x<\theta_2},\quad F(x)=\begin{cases}0, & x<\theta_1;\\ \frac{x-\theta_1}{\theta_2-\theta_1}, & \theta_1\le x<\theta_2;\\ 1, & x\ge\theta_2.\end{cases}$$




则$(X_{(1)},X_{(n)})$的联合密度函数为

$$p_{1n}(x_{(1)},x_{(n)})=n(n-1)[F(x_{(n)})-F(x_{(1)})]^{n-2}p(x_{(1)})p(x_{(n)})\cdot\mathrm{I}_{x_{(1)}\le x_{(n)}}$$

$$=\frac{n(n-1)(x_{(n)}-x_{(1)})^{n-2}}{(\theta_2-\theta_1)^n}\mathrm{I}_{\theta_1<x_{(1)}\le x_{(n)}<\theta_2},$$

（1）由增补变量法得$U=X_{(n)}-X_{(1)}$的密度函数，

当$0<u<\theta_2-\theta_1$时，

$$p_U(u)=\int_{\theta_1}^{\theta_2-u}\frac{n(n-1)[(u+x_{(1)})-x_{(1)}]^{n-2}}{(\theta_2-\theta_1)^n}dx_{(1)}=\frac{n(n-1)u^{n-2}(\theta_2-\theta_1-u)}{(\theta_2-\theta_1)^n},$$

当$u\le 0$或$u\ge\theta_2-\theta_1$时，$p_U(u)=0$，

令$Y=\frac{U}{\theta_2-\theta_1}=\frac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}$，$Y$的密度函数与分布函数分别为

$$p_Y(y)=(\theta_2-\theta_1)p_U((\theta_2-\theta_1)y)=\begin{cases}n(n-1)y^{n-2}(1-y), & 0<y<1;\\ 0, & \text{其他}.\end{cases}$$

$$F_Y(y)=\begin{cases}0, & y<0;\\ ny^{n-1}-(n-1)y^n, & 0\le y<1;\\ 1, & y\ge 1.\end{cases}$$

可得 $Y$ 服从贝塔分布 $Be(n-1, 2)$，其分布与未知参数 $\theta_1, \theta_2$ 无关，$Y$ 为枢轴量，

其 $p$ 分位数 $y_p = Be_p(n-1, 2)$满足方程 $F_Y(y_p)=ny_p^{n-1}-(n-1)y_p^n=p$，

选取枢轴量 $Y=\dfrac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}$，置信度为 $1-\alpha$，即

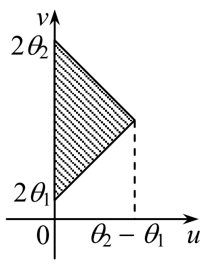
$$P\left\{Be_{\alpha/2}(n-1,2)\le\frac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}\le Be_{1-\alpha/2}(n-1,2)\right\}=1-\alpha,$$

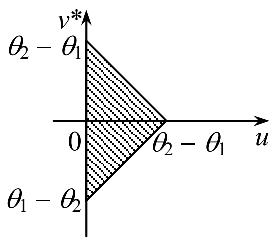
则 $Be_{\alpha/2}(n-1,2)\le\dfrac{X_{(n)}-X_{(1)}}{\theta_2-\theta_1}\le Be_{1-\alpha/2}(n-1,2)$，即 $\dfrac{X_{(n)}-X_{(1)}}{Be_{1-\alpha/2}(n-1,2)}\le\theta_2-\theta_1\le\dfrac{X_{(n)}-X_{(1)}}{Be_{\alpha/2}(n-1,2)}$，

故$\theta_2-\theta_1$ 的置信水平为 $1-\alpha$ 的置信区间为$\left[\dfrac{X_{(n)}-X_{(1)}}{Be_{1-\alpha/2}(n-1,2)},\dfrac{X_{(n)}-X_{(1)}}{Be_{\alpha/2}(n-1,2)}\right]$；

（2）由变量替换公式得$(U, V) = (X_{(n)}-X_{(1)}, X_{(n)}+X_{(1)})$的联合密度函数，有 $X_{(1)}=\dfrac{V-U}{2}, X_{(n)}=\dfrac{V+U}{2}$，

雅可比行列式为 $J=\begin{vmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{vmatrix}=-\dfrac{1}{2}$，




根据$\theta_1<x_{(1)}<x_{(n)}<\theta_2$，可得 $2\theta_1<v-u<v+u<2\theta_2$，
即 $0<u<\theta_2-\theta_1$，$2\theta_1+u<v<2\theta_2-u$，有

$$p_{UV}(u,v)=p_{1n}\left(\frac{v-u}{2},\frac{v+u}{2}\right)\cdot|J|=\frac{n(n-1)u^{n-2}}{2(\theta_2-\theta_1)^n}\cdot \mathrm{I}_{0<u<\theta_2-\theta_1,\ 2\theta_1+u<v<2\theta_2-u},$$

令 $V^* = V-(\theta_2+\theta_1)$，有$(U, V^*)$的联合密度函数为

$$p_{UV^*}(u,v^*)=p_{UV}(u,v^*+(\theta_2+\theta_1))=\frac{n(n-1)u^{n-2}}{2(\theta_2-\theta_1)^n}\cdot \mathrm{I}_{0<u<\theta_2-\theta_1,\ u-(\theta_2-\theta_1)<v<(\theta_2-\theta_1)-u},$$

由增补变量法得 $Z=\dfrac{V^*}{2U}=\dfrac{(X_{(n)}+X_{(1)})-(\theta_2+\theta_1)}{2(X_{(n)}-X_{(1)})}$ 的密度函数，

当 $z<0$ 时，$p_Z(z)=\displaystyle\int_0^{\frac{\theta_2-\theta_1}{1-2z}}\frac{n(n-1)u^{n-2}}{2(\theta_2-\theta_1)^n}\cdot 2u\cdot du=\left.\frac{(n-1)u^n}{(\theta_2-\theta_1)^n}\right|_0^{\frac{\theta_2-\theta_1}{1-2z}}=\frac{n-1}{(1-2z)^n}$，

当 $z\ge 0$ 时，$p_Z(z)=\displaystyle\int_0^{\frac{\theta_2-\theta_1}{1+2z}}\frac{n(n-1)u^{n-2}}{2(\theta_2-\theta_1)^n}\cdot 2u\cdot du=\left.\frac{(n-1)u^n}{(\theta_2-\theta_1)^n}\right|_0^{\frac{\theta_2-\theta_1}{1+2z}}=\frac{n-1}{(1+2z)^n}$，

则 $Z$ 的分布函数为

$$F_Z(z)=\begin{cases}\dfrac{1}{2}(1-2z)^{1-n}, & z<0;\\ 1-\dfrac{1}{2}(1+2z)^{1-n}, & z\ge 0.\end{cases}$$

分布与未知参数$\theta_1, \theta_2$无关，$Z$ 为枢轴量，

当 $p<0.5$ 时，其 $p$ 分位数 $z_p$ 满足 $F_Z(z_p)=\dfrac{1}{2}(1-2z_p)^{1-n}=p$，即 $z_p=\dfrac{1-(2p)^{\frac{1}{1-n}}}{2}$，

当 $p\ge 0.5$ 时，其 $p$ 分位数 $z_p$ 满足 $F_Z(z_p)=1-\dfrac{1}{2}(1+2z_p)^{1-n}=p$，即 $z_p=\dfrac{[2(1-p)]^{\frac{1}{1-n}}-1}{2}$，

选取枢轴量 $Z=\dfrac{(X_{(n)}+X_{(1)})-(\theta_2+\theta_1)}{2(X_{(n)}-X_{(1)})}$，置信度为 $1-\alpha$，即

$$P\left\{z_{\alpha/2}\le\frac{(X_{(n)}+X_{(1)})-(\theta_2+\theta_1)}{2(X_{(n)}-X_{(1)})}\le z_{1-\alpha/2}\right\}=1-\alpha,$$

则 $z_{\alpha/2}=-\dfrac{\alpha^{\frac{1}{1-n}}-1}{2}\le\dfrac{(X_{(n)}+X_{(1)})-(\theta_2+\theta_1)}{2(X_{(n)}-X_{(1)})}\le z_{1-\alpha/2}=\dfrac{\alpha^{\frac{1}{1-n}}-1}{2}$，

即 $\dfrac{X_{(n)}+X_{(1)}}{2}-\dfrac{\alpha^{\frac{1}{1-n}}-1}{2}(X_{(n)}-X_{(1)})\le\dfrac{\theta_2+\theta_1}{2}\le\dfrac{X_{(n)}+X_{(1)}}{2}+\dfrac{\alpha^{\frac{1}{1-n}}-1}{2}(X_{(n)}-X_{(1)})$，

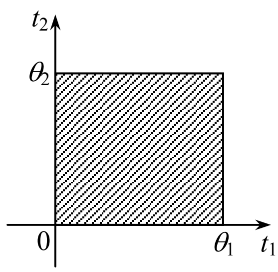
故 $\dfrac{\theta_2+\theta_1}{2}$ 的置信水平为 $1-\alpha$ 的置信区间为

$$\left[\frac{X_{(n)}+X_{(1)}}{2}-\frac{\alpha^{\frac{1}{1-n}}-1}{2}(X_{(n)}-X_{(1)}),\frac{X_{(n)}+X_{(1)}}{2}+\frac{\alpha^{\frac{1}{1-n}}-1}{2}(X_{(n)}-X_{(1)})\right].$$

18．设 $X_1,\cdots,X_m$ i.i.d. $\sim U(0,\theta_1)$，$Y_1,\cdots,Y_n$ i.i.d. $\sim U(0,\theta_2)$，$\theta_1>0$，$\theta_2>0$ 皆未知，且两样本独立，求 $\dfrac{\theta_1}{\theta_2}$ 的一个置信水平为 $1-\alpha$ 的置信区间（提示：令 $T_1=X_{(m)}$，$T_2=Y_{(n)}$，证明 $\dfrac{T_2}{T_1}\cdot\dfrac{\theta_1}{\theta_2}$ 的分布与$\theta_1,\theta_2$无关，并求出对应的密度函数）

证：令 $T_1=X_{(m)}$，$T_2=Y_{(n)}$，有 $T_1$ 与 $T_2$ 相互独立，其联合密度函数为

$$p(t_1,t_2)=p_m(t_1)p_n(t_2)=\frac{mt_1^{m-1}}{\theta_1^m}\mathrm{I}_{0<t_1<\theta_1}\cdot\frac{nt_2^{n-1}}{\theta_2^n}\mathrm{I}_{0<t_2<\theta_2}=\frac{mnt_1^{m-1}t_2^{n-1}}{\theta_1^m\theta_2^n}\mathrm{I}_{0<t_1<\theta_1,\,0<t_2<\theta_2},$$




由增补变量法得 $U=\dfrac{T_2}{T_1}$ 的密度函数，

当 $0<u<\dfrac{\theta_2}{\theta_1}$ 时，

$$p_U(u)=\int_0^{\theta_1}\frac{mnt_1^{m-1}(ut_1)^{n-1}}{\theta_1^m\theta_2^n}\cdot t_1\cdot dt_1=\frac{mnu^{n-1}}{\theta_1^m\theta_2^n}\cdot\frac{t_1^{m+n}}{m+n}\Bigg|_0^{\theta_1}=\frac{mn}{m+n}u^{n-1}\cdot\left(\frac{\theta_1}{\theta_2}\right)^n,$$

当 $u\ge\dfrac{\theta_2}{\theta_1}$ 时，

$$p_U(u)=\int_0^{\frac{\theta_2}{u}}\frac{mnt_1^{m-1}(ut_1)^{n-1}}{\theta_1^m\theta_2^n}\cdot t_1\cdot dt_1=\frac{mnu^{n-1}}{\theta_1^m\theta_2^n}\cdot\frac{t_1^{m+n}}{m+n}\Bigg|_0^{\frac{\theta_2}{u}}=\frac{mn}{m+n}u^{-m-1}\cdot\left(\frac{\theta_2}{\theta_1}\right)^m,$$

当 $u\le 0$ 时，$p_U(u)=0$，

令 $Y=U\cdot\dfrac{\theta_1}{\theta_2}=\dfrac{T_2}{T_1}\cdot\dfrac{\theta_1}{\theta_2}=\dfrac{Y_{(n)}}{X_{(m)}}\cdot\dfrac{\theta_1}{\theta_2}$，$Y$ 的密度函数与分布函数分别为

$$p_Y(y)=\frac{\theta_2}{\theta_1}p_U\left(\frac{\theta_2}{\theta_1}y\right)=\begin{cases}0, & y\le 0;\\ \dfrac{mn}{m+n}y^{n-1}, & 0<y<1;\\ \dfrac{mn}{m+n}y^{-m-1}, & y\ge 1.\end{cases}\quad F_Y(y)=\begin{cases}0, & y<0;\\ \dfrac{m}{m+n}y^n, & 0\le y<1;\\ 1-\dfrac{n}{m+n}y^{-m}, & y\ge 1.\end{cases}$$

分布与未知参数 $\theta_1,\theta_2$ 无关，$Y$ 为枢轴量，

当 $p<\dfrac{m}{m+n}$ 时，其 $p$ 分位数 $y_p$ 满足 $F_Y(y_p)=\dfrac{m}{m+n}y_p^n=p$，即 $y_p=\left[\dfrac{(m+n)p}{m}\right]^{\frac{1}{n}}$，

当 $p\ge\dfrac{m}{m+n}$ 时，其 $p$ 分位数 $y_p$ 满足 $F_Y(y_p)=1-\dfrac{n}{m+n}y_p^{-m}=p$，即 $z_p=\left[\dfrac{n}{(m+n)(1-p)}\right]^{\frac{1}{m}}$，

选取枢轴量 $Y=\dfrac{Y_{(n)}}{X_{(m)}}\cdot\dfrac{\theta_1}{\theta_2}$，置信度为 $1-\alpha$，即 $P\left\{y_{\alpha/2}\le\dfrac{Y_{(n)}}{X_{(m)}}\cdot\dfrac{\theta_1}{\theta_2}\le y_{1-\alpha/2}\right\}=1-\alpha$，

则 $y_{\alpha/2}=\left[\dfrac{(m+n)\alpha}{2m}\right]^{\frac{1}{n}}\le\dfrac{Y_{(n)}}{X_{(m)}}\cdot\dfrac{\theta_1}{\theta_2}\le y_{1-\alpha/2}=\left[\dfrac{2n}{(m+n)\alpha}\right]^{\frac{1}{m}}$，

即 $\dfrac{X_{(m)}}{Y_{(n)}}\left[\dfrac{(m+n)\alpha}{2m}\right]^{\frac{1}{n}}\le\dfrac{\theta_1}{\theta_2}\le\dfrac{X_{(m)}}{Y_{(n)}}\left[\dfrac{2n}{(m+n)\alpha}\right]^{\frac{1}{m}}$，

故 $\dfrac{\theta_1}{\theta_2}$ 的置信水平为 $1-\alpha$ 的置信区间为 $\left[\dfrac{X_{(m)}}{Y_{(n)}}\left[\dfrac{(m+n)\alpha}{2m}\right]^{\frac{1}{n}},\dfrac{X_{(m)}}{Y_{(n)}}\left[\dfrac{2n}{(m+n)\alpha}\right]^{\frac{1}{m}}\right]$.

19．设总体 $X$ 的密度函数为

$$p(x;\theta)=\mathrm{e}^{-(x-\theta)}\mathrm{I}_{x>\theta},\quad -\infty<\theta<\infty$$

$X_1,\cdots,X_n$ 为抽自此总体的简单随机样本.

（1）证明：$X_{(1)}-\theta$ 的分布与 $\theta$ 无关，并求出此分布；

（2）求 $\theta$ 的置信水平为 $1-\alpha$ 的置信区间.

解：（1）总体 $X$ 的分布函数为

$$F(x;\theta)=[1-\mathrm{e}^{-(x-\theta)}]\cdot \mathrm{I}_{x>\theta},$$

则 $X_{(1)}$的密度函数为

$$p_1(x)=n[1-F(x)]^{n-1}p(x)=n\,\mathrm{e}^{-n(x-\theta)}\,\mathrm{I}_{x>\theta},$$

可得 $Y=X_{(1)}-\theta$ 的密度函数为

$$p_Y(y)=p_1(y+\theta)=n\,\mathrm{e}^{-ny}\,\mathrm{I}_{y>0},$$

故 $Y=X_{(1)}-\theta$ 的分布与$\theta$无关，服从指数分布 $Exp(n)$；

（2）因 $Y=X_{(1)}-\theta$ 的分布函数为

$$F_Y(y)=(1-\mathrm{e}^{-ny})\mathrm{I}_{y>0},$$

其 $p$ 分位数 $y_p$ 满足 $F_Y(y_p)=1-\mathrm{e}^{-ny_p}=p$，即 $y_p=-\dfrac{1}{n}\ln(1-p)$，

选取枢轴量 $Y=X_{(1)}-\theta$，置信度为 $1-\alpha$，即 $P\{y_{\alpha/2}\le X_{(1)}-\theta\le y_{1-\alpha/2}\}=1-\alpha$，

则 $y_{\alpha/2}=-\dfrac{1}{n}\ln\left(1-\dfrac{\alpha}{2}\right)\le X_{(1)}-\theta\le y_{1-\alpha/2}=-\dfrac{1}{n}\ln\dfrac{\alpha}{2}$，即 $X_{(1)}+\dfrac{1}{n}\ln\dfrac{\alpha}{2}\le\theta\le X_{(1)}+\dfrac{1}{n}\ln\left(1-\dfrac{\alpha}{2}\right)$，

故$\theta$的置信水平为 $1-\alpha$ 的置信区间为$\left[X_{(1)}+\dfrac{1}{n}\ln\dfrac{\alpha}{2},X_{(1)}+\dfrac{1}{n}\ln\left(1-\dfrac{\alpha}{2}\right)\right]$.