统计学基础 I : 数理统计 Assignment 6

姓名: 雍崔扬
学号: 21307140051
习题: E2.12, E2.16, E2.20, E2.22, E2.23(1)(2)

Problem 1 (习题 2.12)

设 $X=(X_1,\dots,X_n)$ 是取自指数分布 $\{p(x;\mu)=e^{-(x-\mu)}I_{[\mu,\infty)}(x):\mu\in\mathbb{R}\}$ 的简单随机样本.
(1) 求 μ 的最大似然估计量 $\hat{\mu}_1$，并在其基础上得到无偏估计量 $\hat{\mu}_1^\star$

- **Lemma:**
  - $\Gamma(n)=\int_0^\infty x^{n-1}e^{-x}\mathrm{d}x=\begin{cases}1 & \text{if } n=1\\ n-1 & \text{if } n=2,3,\dots\end{cases}$
    $\Gamma(n+1)=n\Gamma(n)$
  - $\Gamma(1)=\int_0^\infty e^{-x}\mathrm{d}x=-e^{-x}|_0^\infty=1$
  - $\Gamma(2)=\int_0^\infty xe^{-x}\mathrm{d}x=-(x+1)e^{-x}|_0^\infty=1$
  - $\Gamma(3)=\int_0^\infty x^2e^{-x}\mathrm{d}x=-(x^2+2x+2)e^{-x}|_0^\infty=2$
- **Solution:**
  似然函数为:

$$
\begin{aligned}
L(\mu|x)&=\prod_{i=1}^n p(x_i;\mu)\\
&=\exp\left\{-\sum_{i=1}^n x_i+n\mu\right\}I\left(\min_{i=1,\dots,n}x_i\geq\mu\right)
\end{aligned}
$$

  容易直接验证 $\hat{\mu}(x)=\min_{i=1,\dots,n}x_i$ 是 $L(\mu|x)$ 的最大值点,
  因此 μ 的 MLE 为 $\hat{\mu}_1=X_{(1)}$
  分布函数 $F(x;\mu)=\int_\mu^x e^{-(t-\mu)}\mathrm{d}t=-e^{-s}|_0^{x-\mu}=1-e^{-(x-\mu)}$

$$
\begin{aligned}
\mathbb{E}[\hat{\mu}_1]&=\mathbb{E}[X_{(1)}]\\
&=\int_\mu^\infty x\cdot n(1-F(x;\mu))^{n-1}p(x;\mu)\mathrm{d}x\\
&=\int_\mu^\infty x\cdot ne^{-(n-1)(x-\mu)}e^{-(x-\mu)}\mathrm{d}x\\
&=n\int_\mu^\infty xe^{-n(x-\mu)}\mathrm{d}x\\
&=n\left[\frac{1}{n^2}\int_\mu^\infty n(x-\mu)e^{-n(x-\mu)}\mathrm{d}(n(x-\mu))+\frac{\mu}{n}\int_\mu^\infty e^{-n(x-\mu)}\mathrm{d}(n(x-\mu))\right]\\
&=\frac{1}{n}\int_0^\infty se^{-s}\mathrm{d}s+\mu\int_0^\infty e^{-s}\mathrm{d}s\\
&=\frac{1}{n}\cdot\Gamma(2)+\mu\cdot\Gamma(1)\\
&=\frac{1}{n}\cdot 1+\mu\cdot 1\\
&=\frac{1}{n}+\mu
\end{aligned}
$$

因此 $\hat{\mu}_1 = X_{(1)}$ 不是 μ 的无偏估计量，
我们可以构造 $\hat{\mu}_1^\star = X_{(n)} - \frac{1}{n}$ 作为 μ 的无偏估计量.

(2) 证明 μ 的矩估计量 $\hat{\mu}_2$ 是 μ 的无偏估计量.

- **Solution:**

$$\begin{aligned}
\mathbb{E}[\xi] &= \int_\mu^\infty x \cdot e^{-(x-\mu)}\mathrm{d}x \\
&= \int_\mu^\infty (x-\mu)e^{-(x-\mu)}\mathrm{d}(x-\mu) + \mu\int_\mu^\infty e^{-(x-\mu)}\mathrm{d}(x-\mu) \\
&= \int_0^\infty se^{-s}\mathrm{d}s + \mu\int_0^\infty e^{-s}\mathrm{d}s \\
&= \Gamma(2) + \mu\cdot\Gamma(1) \\
&= 1 + \mu\cdot 1 \\
&= 1+\mu
\end{aligned}$$

因此有 $\overline{X} = 1 + \hat{\mu}_2$
得到矩估计量 $\hat{\mu}_2 = \overline{X} - 1$

$$\begin{aligned}
\mathbb{E}[\hat{\mu}_2] &= \mathbb{E}[\overline{X}] - 1 \\
&= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n X_i\right] - 1 \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] - 1 \\
&= \frac{1}{n}\cdot n(\mu+1) - 1 \\
&= \mu
\end{aligned}$$

因此矩估计量 $\hat{\mu}_2 = \overline{X} - 1$ 是 μ 的无偏估计量.

(3) $\hat{\mu}_1^\star$ 和 $\hat{\mu}_2$ 哪个更有效?

- **Solution:**
根据 (1) 可知 $\mathbb{E}[X_{(1)}] = \mu + \frac{1}{n}$

$$\begin{aligned}
\mathbb{E}[X_{(1)}^2] &= \int_\mu^\infty x^2 \cdot n(1-F(x;\mu))^{n-1}p(x;\mu)\mathrm{d}x \\
&= n\int_\mu^\infty x^2 e^{-n(x-\mu)}\mathrm{d}x \\
&= n\int_0^\infty \left(\frac{s^2}{n^2} + \frac{2\mu s}{n} + \mu^2\right)e^{-s}\cdot\frac{1}{n}\mathrm{d}s \quad (s := n(x-\mu)) \\
&= \frac{1}{n^2}\int_0^\infty s^2e^{-s}\mathrm{d}s + \frac{2\mu}{n}\int_0^\infty se^{-s}\mathrm{d}s + \mu^2\int_0^\infty e^{-s}\mathrm{d}s \\
&= \frac{1}{n^2}\cdot\Gamma(3) + \frac{2\mu}{n}\cdot\Gamma(2) + \mu^2\cdot\Gamma(1) \\
&= \frac{1}{n^2}\cdot 2 + \frac{2\mu}{n}\cdot 1 + \mu^2\cdot 1 \\
&= \frac{2}{n^2} + \frac{2\mu}{n} + \mu^2
\end{aligned}$$

因此我们有:

$$\begin{aligned}\mathrm{Var}(\hat{\mu}_1^\star) &= \mathrm{Var}(X_{(1)} - \frac{1}{n}) \\ &= \mathrm{Var}(X_{(1)}) \\ &= \mathbb{E}[X_{(1)}^2] - (\mathbb{E}[X_{(1)}])^2 \\ &= \frac{2}{n^2} + \frac{2\mu}{n} + \mu^2 - \left(\mu + \frac{1}{n}\right)^2 \\ &= \frac{1}{n^2}\end{aligned}$$

根据 (2) 可知 $\mathbb{E}[\xi] = \mu + 1$

$$\begin{aligned}\mathbb{E}[\xi^2] &= \int_\mu^\infty x^2 \cdot e^{-(x-\mu)}\mathrm{d}x \\ &= \int_0^\infty (s^2 + 2\mu s + \mu^2)e^{-s}\mathrm{d}s \quad (s := x - \mu) \\ &= \int_0^\infty s^2 e^{-s}\mathrm{d}s + 2\mu\int_0^\infty s e^{-s}\mathrm{d}s + \mu^2\int_0^\infty e^{-s}\mathrm{d}s \\ &= \Gamma(3) + 2\mu\cdot\Gamma(2) + \mu^2\cdot\Gamma(1) \\ &= 2 + 2\mu\cdot 1 + \mu^2\cdot 1 \\ &= 2 + 2\mu + \mu^2\end{aligned}$$

因此我们有:

$$\begin{aligned}\mathrm{Var}(\hat{\mu}_2) &= \mathrm{Var}(\overline{X} - 1) \\ &= \mathrm{Var}(\overline{X}) \\ &= \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) \\ &= \frac{1}{n^2}\cdot n\{\mathbb{E}[\xi^2] - (\mathbb{E}[\xi])^2\} \\ &= \frac{1}{n}[2 + 2\mu + \mu^2 - (\mu+1)^2] \\ &= \frac{1}{n}\end{aligned}$$

对比 $\begin{cases}\mathrm{Var}(\hat{\mu}_1^\star) = \frac{1}{n^2} \\ \mathrm{Var}(\hat{\mu}_2) = \frac{1}{n}\end{cases}$ 可知 $\hat{\mu}_1^\star$ 是更有效的无偏估计量.

Problem 2 (习题 2.16)

设 $X = (X_1, \ldots, X_n)$ 为取自均匀分布族 $\{\mathrm{Uniform}(\theta - \frac{1}{2}, \theta + \frac{1}{2}) : \theta \in \mathbb{R}\}$ 的简单随机样本.
试证明: 对于任意 $\lambda \in [0, 1]$, $\lambda(X_{(1)} + \frac{1}{2}) + (1-\lambda)(X_{(n)} - \frac{1}{2})$ 都是 θ 的最大似然估计量.

Solution:
似然函数为:

$$
\begin{aligned}
L(\theta|x) &= \prod_{i=1}^n \mathrm{P}\{\text{Uniform}(\theta-\frac{1}{2},\theta+\frac{1}{2})=x_i\} \\
&= \prod_{i=1}^n 1\cdot I(\theta-\frac{1}{2}\le x_i\le\theta+\frac{1}{2}) \\
&= I(\theta-\frac{1}{2}\le \min_{i=1,...,n} x_i) I(\max_{i=1,...,n} x_i \le \theta+\frac{1}{2}) \\
&= I(\theta \le \min_{i=1,...,n} x_i+\frac{1}{2}) I(\max_{i=1,...,n} x_i -\frac{1}{2}\le \theta) \\
&= I(\max_{i=1,...,n} x_i -\frac{1}{2}\le \theta \le \min_{i=1,...,n} x_i+\frac{1}{2})
\end{aligned}
$$

显然对于任意 $\lambda\in[0,1]$，
$\hat{\theta}(x)=\lambda(\min_{i=1,...,n} x_i+\frac{1}{2})+(1-\lambda)(\max_{i=1,...,n} x_i-\frac{1}{2})$ 都能使 $L(\theta|x)$ 取到最大值 1
因此对于任意 $\lambda\in[0,1]$，$\hat{\theta}=\lambda(X_{(1)}+\frac{1}{2})+(1-\lambda)(X_{(n)}-\frac{1}{2})$ 都是 θ 的最大似然估计量.

Problem 3 (习题 2.20)

(1) 设随机变量 X,Y 都是正态分布的，且 $\text{Var}(X)\le\text{Var}(Y)$
试证明: 对于任意 $a>0$ 都有 $\mathrm{P}\{|X-\mathbb{E}[X]|\le a\}\ge\mathrm{P}\{|Y-\mathbb{E}[Y]|\le a\}$ 成立.

- **Solution:**
  对于任意 $a>0$，我们都有:

$$
\begin{aligned}
&\mathrm{P}\{|X-\mathbb{E}[X]|\le a\}-\mathrm{P}\{|Y-\mathbb{E}[Y]|\le a\} \\
&=\mathrm{P}\left\{\left|\frac{X-\mathbb{E}[X]}{\sqrt{\text{Var}(X)}}\right|\le\frac{a}{\sqrt{\text{Var}(X)}}\right\}-\mathrm{P}\left\{\left|\frac{Y-\mathbb{E}[Y]}{\sqrt{\text{Var}(Y)}}\right|\le\frac{a}{\sqrt{\text{Var}(Y)}}\right\} \\
&=\left[1-2\left(1-\Phi\left(\frac{a}{\sqrt{\text{Var}(X)}}\right)\right)\right]-\left[1-2\left(1-\Phi\left(\frac{a}{\sqrt{\text{Var}(Y)}}\right)\right)\right] \\
&=2\left[\Phi\left(\frac{a}{\sqrt{\text{Var}(X)}}\right)-\Phi\left(\frac{a}{\sqrt{\text{Var}(Y)}}\right)\right] \quad (\text{note that } \text{Var}(X)\le\text{Var}(Y)) \\
&\ge 0
\end{aligned}
$$

  其中 $\Phi(\cdot)$ 是标准正态分布的密度函数.
  命题得证.

(2) 设随机变量 X,Y 满足 $\mathrm{P}\{|X-a|\le t\}\ge\mathrm{P}\{|Y-a|\le t\}\ \ (\forall\, t>0)$
试证明: $\mathbb{E}[|X-a|^2]\le\mathbb{E}[|Y-a|^2]$

- **引理 1: (Tonelli 定理, Fubini 定理的一个特殊情况)**
  设 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$ 是两个测度空间，
  $f:X\times Y\to[0,\infty]$ 是一个**非负** $\mathcal{A}\times\mathcal{B}$-可测函数.
  则有:
  - 对于任意给定的 $x\in X$，函数 $y\mapsto f(x,y)$ 都是 $\mathcal{B}$-可测的.
  - 对于任意给定的 $y\in Y$，函数 $x\mapsto f(x,y)$ 都是 $\mathcal{A}$-可测的.
  - 函数 $x\mapsto\int_Y f(x,y)\mathrm{d}\nu(y)$ 是 $\mathcal{A}$-可测的，函数 $y\mapsto\int_X f(x,y)\mathrm{d}\mu(x)$ 是 $\mathcal{B}$-可测的.
  - 两个迭代积分是相等的 (即积分可交换顺序):

$$
\int_X\left(\int_Y f(x,y)\mathrm{d}\nu(y)\right)\mathrm{d}\mu(x)=\int_Y\left(\int_X f(x,y)\mathrm{d}\mu(x)\right)\mathrm{d}\nu(y)
$$

- **引理 2: (非负随机向量的原点矩)**
  对于任意**非负**随机变量 X (无论是离散的、连续的，还是这两者的混合形式)
  我们都有 $X = \int_0^X 1\mathrm{d}t = \int_0^\infty \mathbf{1}_{\{X>t\}}(t)\mathrm{d}t$
  应用**引理** 1 **(Tonelli 定理)** 得到:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{X>t\}}(t)\mathrm{d}t\right] \\ &= \int_0^\infty \mathbb{E}[\mathbf{1}_{\{X>t\}}(t)]\mathrm{d}t \\ &= \int_0^\infty \mathrm{P}\{X>t\}\mathrm{d}t\end{aligned}$$

  (由于指示函数 $\mathbf{1}_{\{X>t\}}(t)$ 是非负的，Tonelli 定理允许我们交换期望和积分的顺序)
  类似地，我们有 $X^2 = \int_0^X 2t\mathrm{d}t = \int_0^\infty 2t\cdot\mathbf{1}_{\{X>t\}}(t)\mathrm{d}t$
  应用**引理** 1 **(Tonelli 定理)** 得到:

$$\begin{aligned}\mathbb{E}[X^2] &= \mathbb{E}\left[\int_0^\infty 2t\cdot\mathbf{1}_{\{X>t\}}(t)\mathrm{d}t\right] \\ &= \int_0^\infty 2t\cdot\mathbb{E}[\mathbf{1}_{\{X>t\}}(t)]\mathrm{d}t \\ &= \int_0^\infty 2t\cdot\mathrm{P}\{X>t\}\mathrm{d}t\end{aligned}$$

  (由于 $2t\cdot\mathbf{1}_{\{X>t\}}(t)$ 是非负的，Tonelli 定理允许我们交换期望和积分的顺序)

- **Solution:**
  根据 $\mathrm{P}\{|X-a|\leq t\}\geq\mathrm{P}\{|Y-a|\leq t\}\ \ (\forall\, t>0)$
  我们知道 $\mathrm{P}\{|X-a|>t\}\leq\mathrm{P}\{|Y-a|>t\}\ \ (\forall\, t>0)$

  应用**引理** 2 可知:

$$\begin{aligned}\mathbb{E}[|X-a|^2]-\mathbb{E}[|Y-a|^2] &= \int_0^\infty 2t\cdot\mathrm{P}\{|X-a|>t\}\mathrm{d}t - \int_0^\infty 2t\cdot\mathrm{P}\{|Y-a|>t\}\mathrm{d}t \\ &= \int_0^\infty 2t\cdot(\mathrm{P}\{|X-a|>t\}-\mathrm{P}\{|Y-a|>t\})\mathrm{d}t \\ &\leq 0\end{aligned}$$

  命题得证.

- **Another Solution:**
  根据 $\mathrm{P}\{|X-a|\leq t\}\geq\mathrm{P}\{|Y-a|\leq t\}\ \ (\forall\, t>0)$
  我们知道 $\mathrm{P}\{|X-a|>t\}\leq\mathrm{P}\{|Y-a|>t\}\ \ (\forall\, t>0)$
  因此 $\mathrm{P}\{(X-a)^2>t\}\leq\mathrm{P}\{(Y-a)^2>t\}\ \ (\forall\, t>0)$

  应用**引理** 2 可知:

$$\begin{aligned}\mathbb{E}[|X-a|^2]-\mathbb{E}[|Y-a|^2] &= \int_0^\infty \mathrm{P}\{(X-a)^2>t\}\mathrm{d}t - \int_0^\infty \mathrm{P}\{(Y-a)^2>t\}\mathrm{d}t \\ &= \int_0^\infty (\mathrm{P}\{(X-a)^2>t\}-\mathrm{P}\{(Y-a)^2>t\})\mathrm{d}t \\ &\leq 0\end{aligned}$$

  命题得证.

Problem 4 (习题 2.22)

设总体 ξ 的密度函数为 $p(x;\theta)=\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}I_{(0,\infty)}(x)$，其中 $\theta>0$ 为未知参数.
求 θ 的 Fisher 信息量，以及当样本量为 n 时 θ 无偏估计量方差的 C-R 下界.

Solution:
单个样本的对数似然函数为:

$$\begin{aligned} l(\theta|x) &= \log\left(p(x;\theta)\right) \\ &= \log\left(\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}\right) \\ &= \log\left(2\theta\right) - 3\log\left(x\right) - \frac{\theta}{x^2} \end{aligned}$$

对 θ 求偏导:

$$\frac{\partial}{\partial\theta}l(\theta|x)=\frac{1}{\theta}-\frac{1}{x^2}$$

则 Fisher 信息量 $I_\xi(\theta)=\mathbb{E}_\theta[(\frac{\partial}{\partial\theta}l(\theta|\xi))^2]=\frac{1}{\theta^2}-\frac{1}{\theta}\mathbb{E}[\frac{1}{X^2}]+\mathbb{E}[\frac{1}{X^4}]$

$$\begin{aligned} \mathbb{E}\left[\frac{1}{X^2}\right] &= \int_0^\infty \frac{1}{x^2}\cdot\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}\mathrm{d}x \\ &= \int_0^\infty \frac{2\theta}{x^5}e^{-\frac{\theta}{x^2}}\mathrm{d}x \\ &= -\frac{1}{\theta}\int_\infty^0 ue^{-u}\mathrm{d}u \quad (u=\frac{\theta}{x^2}\Rightarrow \mathrm{d}u=-\frac{2\theta}{x^3}\mathrm{d}x) \\ &= \frac{1}{\theta}\cdot\Gamma(2) \\ &= \frac{1}{\theta}\cdot 1 \\ &= \frac{1}{\theta} \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left[\frac{1}{X^4}\right] &= \int_0^\infty \frac{1}{x^4}\cdot\frac{2\theta}{x^3}e^{-\frac{\theta}{x^2}}\mathrm{d}x \\ &= \int_0^\infty \frac{2\theta}{x^7}e^{-\frac{\theta}{x^2}}\mathrm{d}x \\ &= -\frac{1}{\theta^2}\int_\infty^0 u^2e^{-u}\mathrm{d}u \quad (u=\frac{\theta}{x^2}\Rightarrow \mathrm{d}u=-\frac{2\theta}{x^3}\mathrm{d}x) \\ &= \frac{1}{\theta^2}\cdot\Gamma(3) \\ &= \frac{1}{\theta^2}\cdot 2 \\ &= \frac{2}{\theta^2} \end{aligned}$$

因此 **Fisher 信息量**为:

$$
\begin{aligned}
I_\xi(\theta) &= \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}l(\theta|\xi)\right)^2\right] \\
&= \frac{1}{\theta^2} - \frac{1}{\theta}\mathbb{E}\left[\frac{1}{X^2}\right] + \mathbb{E}\left[\frac{1}{X^4}\right] \\
&= \frac{1}{\theta^2} - \frac{1}{\theta}\cdot\frac{1}{\theta} + \frac{1}{\theta^2} \\
&= \frac{1}{\theta^2}
\end{aligned}
$$

实际上，使用 $I_\xi(\theta) = \mathbb{E}_\theta[(\frac{\partial}{\partial\theta}l(\theta|\xi))^2] = \mathbb{E}_\theta[\frac{\partial^2}{\partial\theta^2}l(\theta|\xi)] = \frac{1}{\theta^2}$ 可以直接得到 Fisher 信息量.
这个式子成立是建立在 "$\int p(x;\theta)\mathrm{d}x = 1$ 关于 θ 可在积分号下微分两次" 的条件下的.

因此当样本量为 n 时 θ 无偏估计量方差的 C-R 下界为:

$$
\frac{(\frac{\mathrm{d}}{\mathrm{d}\theta}\theta)^2}{I_X(\theta)} = \frac{1^2}{nI_\xi(\theta)} = \frac{\theta^2}{n}
$$

也就是说，对于 θ 的无偏估计量 $\hat{\theta}$，都有 $\mathrm{Var}_\theta(\hat{\theta}) \geq \frac{\theta^2}{n}$ 成立.

Problem 5 (习题 2.23 (1)(2))

写出下列分布族中，
达到 C-R 下界的基于样本 $X = (X_1, \ldots, X_n)$ 的无偏估计的参数函数形式和估计量.

(1) 二项分布族 $\{B(k,p) : p \in (0,1)\}$

- **Lemma:**
  若 $p(x;\theta)$ 关于 θ 二阶可偏导，且 $\int p(x;\theta)\mathrm{d}x = 1$ 关于 θ 可在积分号下微分两次，
  则我们有 $I_X(\theta) = \mathbb{E}_\theta[(\frac{\partial}{\partial\theta}l(\theta|X))^2] = -\mathbb{E}_\theta[\frac{\partial^2}{\partial\theta^2}l(\theta|X)]$ 成立.

  证明:

$$
\begin{aligned}
\frac{\partial}{\partial\theta}l(\theta|x) &= \frac{1}{p(x;\theta)}\frac{\partial}{\partial\theta}p(x;\theta) \\
\frac{\partial^2}{\partial\theta^2}l(\theta|x) &= \frac{1}{p(x;\theta)}\frac{\partial^2}{\partial\theta^2}p(x;\theta) - \frac{1}{(p(x;\theta))^2}\left(\frac{\partial}{\partial\theta}p(x;\theta)\right)^2 \\
&= \frac{1}{p(x;\theta)}\frac{\partial^2}{\partial\theta^2}p(x;\theta) - \left(\frac{\partial}{\partial\theta}l(\theta|x)\right)^2
\end{aligned}
$$

  利用 $\int p(x;\theta)\mathrm{d}x = 1$ 关于 θ 可在积分号下微分两次的假设，
  我们有 $\int \frac{\partial^2}{\partial\theta^2}p(x;\theta)\mathrm{d}x = 0$
  即有 $\mathbb{E}_\theta[\frac{1}{p(x;\theta)}\frac{\partial^2}{\partial\theta^2}p(x;\theta)] = \int \frac{1}{p(x;\theta)}\frac{\partial^2}{\partial\theta^2}p(x;\theta)\cdot p(x;\theta)\mathrm{d}x = \int \frac{\partial^2}{\partial\theta^2}p(x;\theta)\mathrm{d}x = 0$

  因此我们有:

$$
\begin{aligned}
\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}l(\theta|x)\right] &= \mathbb{E}_\theta\left[\frac{1}{p(x;\theta)}\frac{\partial^2}{\partial\theta^2}p(x;\theta)\right] - \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}l(\theta|x)\right)^2\right] \\
&= 0 - I_X(\theta)
\end{aligned}
$$

  于是有 $I_X(\theta) = \mathbb{E}_\theta[(\frac{\partial}{\partial\theta}l(\theta|X))^2] = -\mathbb{E}_\theta[\frac{\partial^2}{\partial\theta^2}l(\theta|X)]$ 成立.

Solution:
对数似然函数 $l(p;x)=\log\{\binom{k}{x}p^x(1-p)^{k-x}\}=\log\{\binom{k}{x}\}+x\log{(p)}+(k-x)\log{(1-p)}$
我们有 $\begin{cases}\frac{\partial}{\partial p}l(p;x)=\frac{x}{p}-\frac{k-x}{1-p}\\\frac{\partial^2}{\partial p^2}l(p;x)=-\frac{x}{p^2}-\frac{k-x}{(1-p)^2}\end{cases}$
Fisher 信息量为 (第二步的转化基于 "$\int p(x;p)\mathrm{d}x=1$ 关于 p 可在积分号下微分两次" 的条件):

$$\begin{aligned}I_\xi(p)&=\mathbb{E}_p\left[\left(\frac{\partial}{\partial p}l(p|\xi)\right)^2\right]\\&=-\mathbb{E}_p\left[\frac{\partial^2}{\partial p^2}l(p|\xi)\right]\\&=-\mathbb{E}_p\left[-\frac{\xi}{p^2}-\frac{k-\xi}{(1-p)^2}\right]\\&=\frac{\mathbb{E}[\xi]}{p^2}+\frac{k-\mathbb{E}[\xi]}{(1-p)^2}\\&=\frac{kp}{p^2}+\frac{k-kp}{(1-p)^2}\\&=k\left(\frac{1}{p}+\frac{1}{1-p}\right)\\&=\frac{k}{p(1-p)}\end{aligned}$$

因此 C-R 下界为:

$$\frac{(\frac{\mathrm{d}}{\mathrm{d}p}p)^2}{I_X(p)}=\frac{1^2}{nI_\xi(p)}=\frac{1}{n\cdot\frac{k}{p(1-p)}}=\frac{p(1-p)}{nk}$$

显然矩估计量 $\hat{p}=\frac{1}{k}\overline{X}$ 是 p 的无偏估计量，其方差 $\mathrm{Var}_p(\hat{p})=\frac{p(1-p)}{nk}$ 达到了 C-R 下界，
因此矩估计量 $\hat{p}=\frac{1}{k}\overline{X}$ 是参数 p 的**一致最小方差无偏估计量** UMVUE

---

(2) 正态分布族 $\{N(\mu,\sigma_0^2):\mu\in\mathbb{R}\}$

Solution:
对数似然函数 $l(\mu;x)=-\frac{1}{2}\log\left(2\pi\sigma_0^2\right)-\frac{1}{2\sigma_0^2}(x-\mu)^2$
我们有 $\begin{cases}\frac{\partial}{\partial\mu}l(\mu;x)=\frac{1}{\sigma_0^2}(x-\mu)\\\frac{\partial^2}{\partial\mu^2}l(\mu;x)=-\frac{1}{\sigma_0^2}\end{cases}$

Fisher 信息量为 (第二步的转化基于 "$\int p(x;\mu)\mathrm{d}x=1$ 关于 μ 可在积分号下微分两次" 的条件):

$$\begin{aligned}I_\xi(\mu)&=\mathbb{E}_\mu\left[\left(\frac{\partial}{\partial\mu}l(\mu|\xi)\right)^2\right]\\&=-\mathbb{E}_\mu\left[\frac{\partial^2}{\partial\mu^2}l(\mu|\xi)\right]\\&=-\mathbb{E}_\mu\left[-\frac{1}{\sigma_0^2}\right]\\&=\frac{1}{\sigma_0^2}\end{aligned}$$

因此 C-R 下界为:

$$\frac{(\frac{\mathrm{d}}{\mathrm{d}\mu}\mu)^2}{I_X(\mu)} = \frac{1^2}{nI_\xi(\mu)} = \frac{1}{n \cdot \frac{1}{\sigma_0^2}} = \frac{\sigma_0^2}{n}$$

而样本均值 $\hat{\mu} = \overline{X}$ 作为 μ 的无偏估计量，其方差 $\mathrm{Var}_\mu(\hat{\mu}) = \frac{\sigma_0^2}{n}$ 达到了 C-R 下界，
因此样本均值 $\hat{\mu} = \overline{X}$ 是参数 μ 的**一致最小方差无偏估计量** UMVUE

The End