

第四次作业

题目 1. (22) 设 X_1, \dots, X_n 是来满参数为 θ 的 Bernoulli 分布的随机样本.

- (1). 求解 $\theta(1 - \theta)$ 的 C-R 下界.
- (2). 若 $\theta(1 - \theta)$ 的 UMVUE 存在, 对其进行求解.

解答. (1). Bernoulli 分布的概率密度函数为 $f(x; \theta) = \theta^x(1 - \theta)^{1-x}I_{\{0,1\}}(x)$, 则

$$\frac{\partial}{\partial \theta}(\log \theta^x(1 - \theta^{1-x})) = \frac{x}{\theta} - \frac{1-x}{1-\theta}, \quad \frac{\partial^2}{\partial \theta^2}(\log \theta^x(1 - \theta)^{1-x}) = -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2},$$

则 Fisher 信息量为 $I(\theta) = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2}(\log \theta^x(1 - \theta)^{1-x}) \right] = \frac{1}{(1-\theta)^2} + \left(\frac{1}{\theta^2} - \frac{1}{(1-\theta)^2} \right) \theta = \frac{1}{\theta(1-\theta)}$. 于是 $\theta(1 - \theta)$ 的 C-R 下界为 $\frac{(1-2\theta)^2}{n \frac{1}{\theta(1-\theta)}} = \frac{(1-2\theta)^2 \theta(1-\theta)}{n}$.

(2). 由于 $f(x; \theta) = (1-\theta)I_{\{0,1\}}(x) \exp \left\{ x \log \frac{\theta}{1-\theta} \right\}$, 令 $a(\theta) = 1-\theta, b(x) = I_{\{0,1\}}(x), c(\theta) = \log \frac{\theta}{1-\theta}, d(x) = x$, 则该分布属于指数族, 于是 $\sum_{i=1}^n X_i$ 为完备的充分统计量, 令 $Y = \sum_{i=1}^n X_i, Z = \left(\sum_{i=1}^n X_i \right)^2$, 由于

$$\mathbb{E}(Y) = n\theta, \quad \mathbb{E}(Z) = \mathbb{E} \left[\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j \right] = n\theta + n(n-1)\theta^2 = n(n-1) \left(\frac{\theta}{n-1} + \theta^2 \right).$$

则 $\mathbb{E} \left[\frac{Y}{n} \right] = \theta, \quad \mathbb{E} \left[\frac{Z}{n(n-1)} - \frac{Y}{n(n-1)} \right] = \theta^2$, 于是

$$\theta(1 - \theta) = \mathbb{E} \left[\frac{Y}{n} \right] - \mathbb{E} \left[\frac{Z}{n(n-1)} - \frac{Y}{n(n-1)} \right] = \mathbb{E} \left[\frac{Y}{n-1} - \frac{Z}{n(n-1)} \right],$$

由于 $\frac{1}{n-1} \sum_{i=1}^n X_i - \frac{1}{n(n-1)} \left(\sum_{i=1}^n X_i \right)^2$ 是关于完备充分统计量 Y 的函数, 且是 $\theta(1 - \theta)$ 的无偏估计, 所以它是 $\theta(1 - \theta)$ 的 UMVUE.

题目 2. (24) 设 X_1, \dots, X_n 是来自 $\theta x^{\theta-1} I_{(0,1)}(x)$, ($\theta > 0$) 的随机样本,

- (1). 求解 $\mu = \theta/(1 + \theta)$ 的 MLE.
- (2). 求解一个充分统计量, 并验证其完备性. 判断 $\sum_{i=1}^n X_i$ 是否是充分统计量?
- (3). 是否存在一个关于 θ 的函数, 使得其存在无偏估计且方差满足 C-R 下界?
- (4). 求解 $\theta, 1/\theta, \mu = \theta/(1 + \theta)$ 的 UMVUE.

解答. (1). 只需求解 θ 的 MLE: $\mathcal{L}(\theta) = \prod_{i=1}^n \theta x_i^{\theta-1} I_{(0,1)}(x_i)$, 由

$$\frac{\partial}{\partial \theta} \log \mathcal{L}(\theta) = \frac{\partial}{\partial \theta} \left(n \log \theta + \sum_{i=1}^n (\theta - 1) \log(x_i) + \sum_{i=1}^n \log I_{(0,1)}(x_i) \right) = \frac{n}{\theta} + \sum_{i=1}^n \log(x_i) = 0$$

可得 θ 的 MLE 为 $\hat{\theta} = -\frac{n}{\sum_{i=1}^n \log(x_i)}$. 则 μ 的 MLE 为

$$\hat{\mu} = \frac{\hat{\theta}}{1 - \hat{\theta}} = \frac{n}{n - \sum_{i=1}^n \log(x_i)}.$$

(2). 由于 $f(x; \theta) = \theta x^{\theta-1} I_{(0,1)}(x) = \theta I_{(0,1)}(x) \exp\{(\theta-1) \log x\}$, 令 $a(\theta) = \theta$, $b(x) = I_{(0,1)}(x)$, $c(\theta) = \theta - 1$, $d(x) = \log x$, 于是该分布属于指数族, 则 $\sum_{i=1}^n d(x_i) = \sum_{i=1}^n \log x_i$ 是一个完备的充分统计量. 由于 $\sum_{i=1}^n X_i$ 不满足因子分解定理, 所以不是充分统计量.

(3). 由于 $f_{-\log X}(x) = f_X(e^{-x})e^{-x} = \theta e^{-(\theta-1)x} e^{-x} I_{(0,1)}(e^{-x}) = \theta e^{-\theta x} I_{(0,\infty)}(x)$. 于是 $-\log X \sim \text{Exp}(\theta)$. 令 $T = -\frac{1}{n} \sum_{i=1}^n \log X_i$, 则 $nT \sim \Gamma(n, \theta) \Rightarrow \mathbb{E}(T) = \frac{1}{\theta}$. 则 T 是 $\frac{1}{\theta}$ 的无偏估计, 且

$$\sum_{i=1}^n \frac{\partial}{\partial \theta} \log f(x_i; \theta) = \sum_{i=1}^n \frac{\partial}{\partial \theta} (\log \theta + (\theta-1) \log x) = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = -n \left(T - \frac{1}{\theta} \right)$$

所以 C-R 方程取到等号, 故 $T = -\frac{1}{n} \sum_{i=1}^n \log X_i$ 取到 $\frac{1}{\theta}$ 的 C-R 下界.

(4). 由第三小问可知, $-\frac{1}{n} \sum_{i=1}^n \log X_i$ 是 $\frac{1}{\theta}$ 的 UMVUE.

做待定系数, 令 $\frac{c}{-\sum_{i=1}^n \log X_i}$ 是 θ 的无偏估计, 记 $Y = -\sum_{i=1}^n \log X_i$, 则

$$\mathbb{E} \left[\frac{c}{-\sum_{i=1}^n \log X_i} \right] = \int_0^\infty \frac{c}{y} \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y} dy = \frac{c\theta^n}{\Gamma(n)} \cdot \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{c}{n-1} \theta,$$

所以 $c = n-1$, 故 $\frac{n-1}{-\sum_{i=1}^n \log X_i}$ 是 θ 的 UMVUE.

$\theta(1-\theta)$ 的一个无偏估计为 X_1 , 可以得到 $\mathbb{E} \left[X_1 \middle| -\sum_{i=1}^n \log X_i \right] = \frac{n-1}{s} \int_0^1 \left(1 + \frac{\log x}{s}\right)^{n-2} dx$, 但是该积分不会计算, 就没写了 (应该可以换元后凑 Gamma 分布得到积分结果).

题目 3. (29) 设 X_1, \dots, X_n 是来自 $N(\theta, 1)$ 的随机样本.

- (1). 分别求解关于 $\theta, \theta^2, P(X > 0)$ 的 C-R 下界.
- (2). 当 $n = 1$ 时, 是否存在关于 θ^2 的无偏估计? 若存在, 求解之.
- (3). 是否存在关于 $P(X > 0)$ 的无偏估计? 若存在, 求解之.
- (4). 求解 $P(X > 0)$ 的 MLE.
- (5). 是否存在关于 θ^2 的 UMVUE. 若存在, 求解之.
- (6). 是否存在关于 $P(X > 0)$ 的 UMVUE. 若存在, 求解之.

解答. (1). $N(\theta, 1)$ 的概率密度函数为 $f(x; \theta) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(x-\theta)^2}{2} \right\}$, 则

$$\frac{\partial}{\partial \theta} \log f(x; \theta) = \frac{\partial}{\partial \theta} \left(-\frac{\log 2\pi}{2} - \frac{(x-\theta)^2}{2} \right) = x - \theta,$$

Fisher 信息量为 $I(\theta) = \mathbb{E} \left[\frac{\partial}{\partial \theta} \log f(x; \theta) \right]^2 = \mathbb{E}[x^2 - 2\theta x + \theta^2]$, 由于 $\mathbb{E}(x^2) = \text{Var}(X) + [\mathbb{E}(X)]^2 = 1 + \theta^2$, 于是 $I(x) = 1 + \theta^2 - 2\theta^2 + \theta^2 = 1$.

所以 θ 的 C-R 下界为 $\frac{1}{n}$, θ^2 的 C-R 下界为 $\frac{4\theta^2}{n}$.

由于 $X - \theta \sim N(0, 1)$, 则

$$P(x > 0) = P(x - \theta > -\theta) = 1 - P(x - \theta \leq -\theta) = 1 - \Phi(-\theta) = \Phi(\theta),$$

所以 $P(x > 0)$ 的 C-R 下界为 $\frac{(\Phi'(-\theta))^2}{n} = \frac{f(\theta; 1)^2}{n} = \frac{e^{-\theta^2}}{2\pi n}$.

(2). 由于 $\mathbb{E}(X_1^2) = 1 + \theta^2$, 则 $X_1^2 - 1$ 是 θ^2 的无偏估计.

(3). 由于 $\mathbb{E}(I(X_1 > 0)) = \int_0^\infty f(x; \theta) dx = P(X > 0)$, 则 $I(X_1 > 0)$ 是 $P(X > 0)$ 的无偏估计.

(4). 只需求解 θ 的 MLE:

$$\log \mathcal{L}(\theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^n (x_i - \theta)^2, \quad \frac{\partial}{\partial \theta} \log \mathcal{L}(\theta) = \sum_{i=1}^n (x_i - \theta) = 0$$

则 θ 的 MLE 为 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, 由于 $P(X > 0) = 1 - \Phi(-\theta) = \Phi(\theta)$, 则 $P(X > 0)$ 的 MLE 为 $\Phi(\bar{X})$.

(5). 由于

$$\mathbb{E} \left(\sum_{i=1}^n X_i \right)^2 = \text{Var} \left(\sum_{i=1}^n X_i \right) + \left[\mathbb{E} \left(\sum_{i=1}^n X_i \right) \right]^2 = \sum_{i=1}^n \text{Var}(X_i) + (n\mathbb{E}(X))^2 = n + n^2\theta^2$$

则 θ^2 的 UMVUE 为 $\frac{1}{n^2} \left[\left(\sum_{i=1}^n X_i \right)^2 - n \right]$.

(6). 令 $S = \sum_{i=1}^n X_i$, 由于 $I(X_1 > 0)$ 是 $P(X > 0)$ 的 UMVUE, 则

$$\mathbb{E} \left[I(X_1 > 0) \middle| \sum_{i=1}^n X_i = s \right] = P(X_1 > 0 | \sum_{i=1}^n X_i = s) = \int_0^\infty f(x|s) dx$$

令 $Y = \sum_{i=2}^n X_i$, 则 Y 与 X_1 独立, $f_{X_1, Y}(x, y) = \frac{1}{2\pi\sqrt{n-1}} \exp \left\{ -\frac{(x-\theta)^2}{2} - \frac{(y-(n-1)\theta)^2}{2(n-1)} \right\}$.

变量代换得

$$f_{X_1, S}(x, s) = f_{X, Y}(x, s-x) \cdot 1 = \frac{1}{2\pi\sqrt{n-1}} \exp \left\{ -\frac{(x-\theta)^2}{2} - \frac{(s-x-(n-1)\theta)^2}{2(n-1)} \right\}$$

于是

$$\begin{aligned}
f(x|s) &= \frac{f(x, s)}{f(s)} = \sqrt{\frac{n}{2\pi(n-1)}} \exp \left\{ -\frac{(x-\theta)^2}{2} - \frac{(s-x-(n-1)\theta)^2}{2(n-1)} + \frac{(s-n\theta)^2}{2n} \right\} \\
&= \int_0^\infty \sqrt{\frac{n}{2\pi(n-1)}} \exp \left\{ -\frac{1}{2} \left(\sqrt{\frac{n}{n-1}}x - \frac{s}{\sqrt{n(n-1)}} \right)^2 \right\} dx \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(x - \frac{s}{\sqrt{n(n-1)}} \right)^2 \right\} dx \\
&= 1 - \Phi\left(-\frac{s}{\sqrt{n(n-1)}}\right) = \Phi\left(\frac{s}{\sqrt{n(n-1)}}\right)
\end{aligned}$$

则 $P(X > 0)$ 的 UMVUE 为 $\Phi\left(\frac{n\bar{X}}{\sqrt{n(n-1)}}\right)$.

题目 4. (30) 设 X_1, \dots, X_n 是来自参数为 λ 的 Poisson 分布随机样本, 求解关于 $\tau(\lambda) = (1+\lambda)e^{-\lambda}$ 的无偏估计量, MLE 和 UMVUE.

解答. Poisson 分布的概率密度函数为 $f(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,\dots\}}(x)$.

由于 $\mathbb{E}(I(X_1 = 0) + I(X_1 = 1)) = P(X_1 = 0) + P(X_1 = 1) = (1 + \lambda)e^{-\lambda}$, 所以 $I(X_1 = 0) + I(X_1 = 1)$ 是 $(1 + \lambda)e^{-\lambda}$ 的无偏估计.

$$\log \mathcal{L}(\lambda) = \log \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} I_{\{0,1,\dots\}}(x) = \sum_{i=1}^n x_i \log \lambda - \sum_{i=1}^n \log x_i! - n\lambda + \sum_{i=1}^n \log I_{\{0,1,\dots\}}(x).$$

于是 $\frac{\partial}{\partial \lambda} \log f(x; \lambda) = \frac{1}{\lambda} \sum_{i=1}^n x_i - n = 0 \Rightarrow \hat{\lambda} = \bar{X}$, 则 $\tau(\lambda)$ 的 MLE 为 $(1 + \bar{X})e^{-\bar{X}}$.

由于 $f(x; \lambda) = e^{-\lambda} \frac{I_{\{0,1,\dots\}}(x)}{x!} \exp\{-x \log \lambda\}$, 令 $a(\lambda) = e^{-\lambda}$, $b(x) = \frac{I_{\{0,1,\dots\}}(x)}{x!}$, $c(\lambda) = -\log \lambda$, $d(x) = x$, 则该分布属于指数族, 于是 $\sum_{i=1}^n X_i = n\bar{X}$ 是完备充分统计量. 则

$$\mathbb{E}[I(X_1 = 0) + I(X_1 = 1)|n\bar{X} = s] = P[X_1 = 0|n\bar{X} = s] + P[X_1 = 1|n\bar{X} = s]$$

由于 $n\bar{X} \sim \text{Poi}(n\lambda)$, 于是 $P(n\bar{X} = s) = \frac{(n\lambda)^s}{s!} e^{-n\lambda}$, 且

$$P(X_1 = 0, n\bar{X} = s) = P(X_1 = 0)P\left(\sum_{i=2}^n X_i = s\right) = e^{-\lambda} \cdot \frac{((n-1)\lambda)^s}{s!} e^{-(n-1)\lambda} = \frac{(n-1)^s \lambda^s}{s!} e^{-n\lambda}$$

$$P(X_1 = 1, n\bar{X} = s) = P(X_1 = 1)P\left(\sum_{i=2}^n X_i = s-1\right) = \lambda e^{-\lambda} \cdot \frac{((n-1)\lambda)^{s-1}}{(s-1)!} e^{-(n-1)\lambda} = \frac{(n-1)^{s-1} \lambda^s}{(s-1)!} e^{-n\lambda}$$

所以

$$\begin{aligned}
\mathbb{E}[I(X_1 = 0) + I(X_1 = 1)|n\bar{X} = s] &= \frac{P(X_1 = 0, n\bar{X} = s)}{P(n\bar{X} = s)} + \frac{P(X_1 = 1, n\bar{X} = s)}{P(n\bar{X} = s)} \\
&= \left(\frac{n-1}{n}\right)^s + \frac{s}{n} \left(\frac{n-1}{n}\right)^{s-1}.
\end{aligned}$$

则 $\left(\frac{n-1}{n}\right)^{n\bar{X}} + \bar{X} \left(\frac{n-1}{n}\right)^{n\bar{X}-1}$ 是 $\tau(\lambda)$ 的 UMVUE.

题目 5. (31) 设 X_1, \dots, X_n 是来自 $f(x; \theta) = \frac{2x}{\theta^2} I_{(0, \theta)}(x)$, ($\theta > 0$) 的随机样本.

(1). 求解关于 θ 的 MLE.

(2). $Y_n = \max\{X_1, \dots, X_n\}$ 是充分统计量吗? Y_n 是完备的吗?

(3). 是否存在关于 θ 的 UMVUE. 若存在, 求解之.

解答. (1). 极大似然函数为

$$\mathcal{L}(\theta) = \prod_{i=1}^n \frac{2x_i}{\theta^2} I_{(0, \theta)}(x_i) = \frac{2}{\theta^{2n}} I_{[0, y_n]}(y_1) I_{[y_n, \infty)}(\theta) \prod_{i=1}^n x_i.$$

则 θ 的 MLE 为 $\hat{\theta} = Y_n$.

(2). 由于 $\mathcal{L}(\theta) = I_{[0, y_n]}(y_1) \prod_{i=1}^n x_i \cdot \frac{2}{\theta^{2n}} I_{[y_n, \infty)}(\theta) = h(x_1, \dots, x_n) \cdot g(y_n; \theta)$, 所以 Y_n 是充分统

计量. 由于 $f_{Y_n}(y) = nF(y)^{n-1}f(y) = n \left(\frac{y^2}{\theta^2}\right)^{n-1} \cdot \frac{2y}{\theta^2} = \frac{2n}{\theta^{2n}} y^{2n-1}$, 对于任意的实值函数 $z(\cdot)$, 有下式成立

$$0 = \mathbb{E}(z(Y_n)) = \int_0^\theta z(y) \frac{2n}{\theta^{2n}} y^{2n-1} dy \Rightarrow \int_0^\theta z(y) y^{2n-1} dy = 0$$

对 θ 求导可得 $z(\theta)\theta^{2n-1} = 0$, ($\theta > 0$), 则 $z(\theta) \equiv 0$, 于是 $P(z(Y_n) = 0) = 1$, 故 Y_n 是完备的.

(3). 由于

$$\mathbb{E}(Y_n) = \int_0^\theta y \cdot \frac{2n}{\theta^{2n}} y^{2n-1} dy = \frac{2n}{2n+1} \theta$$

则 $\frac{2n+1}{2n} Y_n$ 是 θ 的 UMVUE.

题目 6. (32) 设 X_1, \dots, X_n 是来自 $f(x; \theta) = \theta(1+x)^{-(1+\theta)} I_{(0, \infty)}(x)$, ($\theta > 0$) 的随机样本.

(1). 当 $\theta > 1$ 时, 求解 θ 的矩估计.

(2). 求解 $1/\theta$ 的 MLE.

(3). 如果完备充分统计量存在, 对其进行求解.

(4). 求解关于 $1/\theta$ 的 C-R 下界.

(5). 求解关于 $1/\theta$ 的 UMVUE.

(6). 求解关于 θ 的 UMVUE.

解答. (1). $\mathbb{E}(\bar{X}) = \mathbb{E}(X) = \int_0^\infty x\theta(1+x)^{-(1+\theta)} dx = \theta \int_1^\infty (x-1)x^{-(1+\theta)} dx = \frac{1}{\theta-1}$, 于是 θ 的矩估计为 $\frac{1}{\bar{X}} + 1$.

(2). 由于 $\log L(\theta) = \log \theta^n \prod_{i=1}^n (1+x_i)^{-(1+\theta)} = n \log \theta - (1+\theta) \sum_{i=1}^n \log(1+x_i)$, 则 $\frac{\partial}{\partial \theta} \log L(\theta) = \frac{n}{\theta} - \sum_{i=1}^n \log(1+x_i) = 0$, 于是 θ 的 MLE 为 $\hat{\theta} = \frac{1}{\frac{1}{n} \sum_{i=1}^n \log(1+x_i)}$, 所以 $1/\theta$ 的 MLE 为 $\frac{n}{\sum_{i=1}^n \log(1+X_i)}$.

(3). 由于 $f(x; \theta) = \theta(1+x)^{-(1+\theta)} I_{(0,\infty)}(x) = \theta I_{(0,\infty)}(x) \exp\{-(1+\theta)\log(1+x)\}$, 令 $a(\theta) = \theta$, $I_{(0,\infty)}(x) = b(x)$, $c(\theta) = -(1+\theta)$, $d(x) = \log(1+x)$, 所以该分布属于指数族, 则 $\sum_{i=1}^n \log(1+X_i)$ 为完备充分统计量.

(4). 由于 $\frac{\partial}{\partial \theta} \log f(x; \theta) = \frac{\partial}{\partial \theta} (\log \theta - (1+\theta) \log(1+x)) = \frac{1}{\theta} - \log(1+x)$, $\frac{\partial^2}{\partial \theta^2} \log f(x; \theta) = -\frac{1}{\theta^2}$, 则 Fisher 信息量为 $I(\theta) = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \log f(x; \theta) \right] = \frac{1}{\theta^2}$, 则 $1/\theta$ 的 C-R 下界为 $\frac{[(1/\theta)']^2}{nI(\theta)} = \frac{1}{n\theta^2}$.

(5). 令 $Y = \log(1+X)$, 则 $f_Y(y) = f_X(e^y - 1)e^y = \theta(e^y)^{-(1+\theta)}e^y = \theta e^{-\theta y} \sim \text{Exp}(\theta)$. 则 $\mathbb{E} \left[\sum_{i=1}^n \log(1+X_i) \right] = n\mathbb{E}(Y) = \frac{n}{\theta}$, 则 $\frac{1}{\theta}$ 的 UMVUE 为 $\frac{1}{n} \sum_{i=1}^n \log(1+X_i)$.

(6). (该过程与 24.(4) 求解 θ 的无偏估计方法完全一致, 均为 Gamma 分布的变式) 做待定系数, 令 $\frac{c}{\sum_{i=1}^n \log(1+X_i)}$ 为 θ 的无偏估计, 记 $Y = \sum_{i=1}^n \log(1+X_i)$, 则 $Y \sim \Gamma(n, \theta)$, 且

$$\mathbb{E} \left[\frac{c}{\sum_{i=1}^n \log(1+X_i)} \right] = \int_0^\infty \frac{c}{y} \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y} dy = \frac{c\theta^n}{\Gamma(n)} \cdot \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{c}{n-1} \theta,$$

所以 θ 的 UMVUE 为 $\frac{n-1}{\sum_{i=1}^n \log(1+X_i)}$.