

第五次作业

题目 1. (44) 令 X_1, \dots, X_n 是来自 $f(x; \theta) = e^{-(x-\theta)} I_{[\theta, \infty)}(x)$, ($\theta \in \mathbb{R}^1$) 的随机样本.

- 求解一个充分统计量.
- 求出 θ 的 MLE.
- 求出 θ 的矩估计.
- 求解一个完备充分统计量.
- 求解 θ 的 UMVUE.
- 求出关于 θ 的 Pitman 估计量.
- 使用先验分布 $g(\theta) = e^{-\theta} I_{(0, \infty)}(\theta)$, 求出 θ 的后验 Bayes 估计量.

解答. (a). $\prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n e^{-(x_i-\theta)} I_{[\theta, \infty)}(x_i) = I_{(-\infty, y_1]}(\theta) e^{n\theta} e^{-\sum_{i=1}^n x_i}$, 则 $Y_1 = \min\{X_1, \dots, X_n\}$ 是一个充分统计量.

(b). 由于 $\mathcal{L}(\theta) = \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n e^{-(x_i-\theta)} I_{[\theta, \infty)}(x_i) = \exp\left\{-\sum_{i=1}^n x_i + n\theta\right\} I_{(-\infty, y_1]}(\theta)$, 其中 $Y_1 = \min\{X_1, \dots, X_n\}$, 则 MLE 为 $\hat{\theta} = Y_1$.

(c). $\mathbb{E}(\bar{X}) = \mathbb{E}(X) = \int_{\theta}^{\infty} x e^{-(x-\theta)} dx = \int_0^{\infty} (x + \theta) e^{-x} dx = \frac{\Gamma(2)}{1^2} + \theta = 1 + \theta$, 于是 θ 的矩估计为 $\tilde{\theta} = \bar{X} - 1$.

(d). 下面证明 Y_1 是完备的, $f_{Y_1} = n[1 - F_x(y)]^{n-1} f_x(y) = n e^{-n(y-\theta)}$, 令 $z(\cdot)$ 是任意实值函数, 满足 $\mathbb{E}(z(Y_1)) = 0$, $\theta \in \mathbb{R}^1$, 于是

$$\mathbb{E}(z(Y_1)) = \int_{\theta}^{\infty} z(y) n e^{-n(y-\theta)} dy = \int_{-\infty}^{-\theta} z(y) e^{-ny} dy = 0$$

对 $-\theta$ 求导可得 $z(-\theta) e^{n\theta} = z(-\theta) = 0$, 由 θ 的任意性可知 $z(\cdot) \equiv 0$, 则 $P(z(Y_1) = 0) = 1$, 所以 Y_1 是完备的, 结合 (a) 可知 Y_1 是完备充分统计量.

(e). 由于 $\mathbb{E}(Y_1) = \int_{\theta}^{\infty} y n e^{-n(y-\theta)} dy = n \int_0^{\infty} (y + \theta) e^{-ny} dy = n \int_0^{\infty} y e^{-ny} dy + \theta \int_0^{\infty} n e^{-ny} dy = \frac{1}{n} + \theta$. 于是 $Y_1 - 1/n$ 是 θ 的 UMVUE.

(g). 由于 $\mathcal{L}(\theta) = e^{n\theta - \sum_{i=1}^n x_i} I_{(-\infty, y_1]}(\theta) \propto e^{n\theta} I_{(-\infty, y_1]}(\theta)$ 且 $f(\theta | x_1, \dots, x_n) = \int \mathcal{L}(\theta) g(\theta) d\theta = \int_0^{y_1} e^{(n-1)\theta} d\theta = \frac{e^{(n-1)y_1} - 1}{n-1}$, ($y_1 > 0$). 则

$$f(\theta | x_1, \dots, x_n) = \frac{\mathcal{L}(\theta) g(\theta)}{\int \mathcal{L}(\theta) g(\theta) d\theta} = \frac{n-1}{e^{(n-1)y_1} - 1} e^{(n-1)\theta} I_{(0, y_1]}(\theta).$$

于是

$$\begin{aligned} \mathbb{E}[\theta | x_1, \dots, x_n] &= \int_0^{y_1} \theta \frac{n-1}{e^{(n-1)y_1} - 1} e^{(n-1)\theta} d\theta = \frac{n-1}{e^{(n-1)y_1} - 1} \int_0^{y_1} \theta e^{-(1-n)\theta} d\theta \\ &= \frac{n-1}{e^{(n-1)y_1} - 1} \cdot \frac{\Gamma(2)}{(1-n)^2} = \frac{1}{(n-1)(e^{(n-1)y_1} - 1)} \end{aligned}$$

则 θ 的后验 Bayes 估计量为 $\frac{1}{(n-1)(e^{(n-1)y_1} - 1)}$.

题目 2. (47) 令 X_1, \dots, X_n 是来自离散分布 $f(x; \theta) = \binom{2}{x} \theta^x (1-\theta)^{2-x} I_{\{0,1,2\}}(x)$, ($\theta > 0$) 的随机变量.

- 是否存在一维的充分统计量, 若存在, 是否完备?
- 求解 $\theta^2 = P(X_1 = 2)$ 的 MLE, 并判断是否是无偏的.
- 求解一个关于 θ 的无偏统计量, 且满足 C-R 下界. 若不存在, 证明之.
- 求解一个 θ^2 的 UMVUE.
- 利用平方损失函数求解 θ 关于先验分布为 Beta 分布的 Bayes 估计. Beta 分布为 $g(\theta) = \frac{1}{B(a, b)} \theta^{a-1} (1-\theta)^{b-1} I_{(0,1)}(\theta)$.
- 使用平方误差损失函数, 求解 θ 的极小化极大估计.
- 求解 θ^2 均方误差的一致估计.

解答. (a). 由于

$$f(x; \theta) = \exp\{2 \log(1-\theta)\} \binom{2}{x} I_{\{0,1,2\}}(x) \exp\left\{x \log \frac{\theta}{1-\theta}\right\}$$

令 $a(\theta) = \exp\{2 \log(1-\theta)\}$, $b(x) = \binom{2}{x} I_{\{0,1,2\}}(x)$, $c(\theta) = \log \frac{\theta}{1-\theta}$, $d(x) = x$. 则 $\sum_{i=1}^n X_i$ 是完备充分统计量.

$$(b). \mathcal{L}(\theta) = \prod_{i=1}^n f(x_i; \theta) = (1-\theta)^{2n} \left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^n x_i} \left[\prod_{i=1}^n \binom{2}{x_i} I_{\{0,1,2\}}(x_i) \right],$$

$$\log \mathcal{L}(\theta) = 2n \log(1-\theta) + \sum_{i=1}^n \frac{\theta}{1-\theta} + \sum_{i=1}^n \log \binom{2}{x_i} I_{\{0,1,2\}}(x_i), \text{ 于是 } \frac{\partial \log \mathcal{L}(\theta)}{\partial \theta} = -\frac{2n}{1-\theta} + \frac{1}{\theta(1-\theta)} \sum_{i=1}^n x_i = 0, \text{ 得到 } \theta \text{ 的 MLE 为 } \hat{\theta} = \frac{\bar{X}}{2}, \text{ 于是 } \hat{\theta}^2 = \frac{\bar{X}^2}{4}.$$

$$\begin{aligned} \mathbb{E}(\bar{X}^2/4) &= \frac{1}{4n^2} \mathbb{E} \left[\sum_{i=1}^n X_i \right]^2 = \frac{1}{4n^2} \mathbb{E} \left[\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j \right] \\ &= \frac{1}{4n^2} [n \mathbb{E}(X^2) + n(n-1)(\mathbb{E}X)^2] = \frac{1}{4n^2} [n(\text{Var}(X) + (\mathbb{E}X)^2) + n(n-1)(\mathbb{E}X)^2] \\ &= \frac{1}{4n^2} [n \text{Var}(X) + n^2 (\mathbb{E}X)^2] = \theta^2 + \frac{\theta(1-\theta)}{2n} \end{aligned}$$

所以 $\hat{\theta}^2$ 不是无偏估计.

(c). 由上一问可知

$$\begin{aligned} \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f(x_i; \theta) &= \frac{\partial \log L(\theta)}{\partial \theta} = -\frac{2n}{1-\theta} + \frac{1}{\theta(1-\theta)} \sum_{i=1}^n x_i \\ &= \frac{2n}{\theta(1-\theta)} \left(\frac{\bar{X}}{2} - \theta \right) \end{aligned}$$

又由于 $\mathbb{E} \left[\frac{\bar{X}}{2} \right] = \mathbb{E}(X)/2 = \theta$, 所以 $\bar{X}/2$ 是 θ 的无偏统计量且满足 C-R 下界.

(d). 由 (b) 小问可知 $\mathbb{E}[\bar{X}^2/4] = \frac{2n-1}{2n}\theta^2 + \frac{\theta}{2n}$, 于是 $\frac{2n}{2n-1} \left(\frac{\bar{X}^2}{4} - \frac{\bar{X}}{2n} \right) = \frac{\bar{X}(n\bar{X}-1)}{2(2n-1)}$

是 θ^2 的无偏估计, 且是关于完备充分统计量 $\sum_{i=1}^n X_i$ 的函数, 于是也是 θ^2 的 UMVUE.

(e). θ 的后验概率密度函数为

$$\begin{aligned} f(\theta|x_1, \dots, x_n) &= \frac{\mathcal{L}(\theta)g(\theta)}{\int_0^1 \mathcal{L}(\theta)g(\theta) d\theta} \propto (1-\theta)^{2n} \left(\frac{\theta}{1-\theta} \right)^{\sum_{i=1}^n x_i} \theta^{a-1} (1-\theta)^{b-1} \\ &\propto \theta^{a-1+\sum_{i=1}^n x_i} (1-\theta)^{b-1+2n-\sum_{i=1}^n x_i} \sim \text{Beta} \left(a + \sum_{i=1}^n x_i, b + 2n - \sum_{i=1}^n x_i \right) \end{aligned}$$

由于平方损失函数下 Bayes 估计就是后验 Bayes 估计, 于是 $\mathbb{E}[\theta|X_1, \dots, X_n] = \frac{a+n\bar{X}}{a+b+2n}$ 是 θ 在先验分布为 Beta 分布时的 Bayes 估计.

(f). 由 (e) 可知, θ 的后验 Bayes 估计为 $\mathbb{E}[\theta|X_1, \dots, X_n] = \frac{a+n\bar{X}}{a+b+2n}$, 由于损失函数为平方损失, 则后验 Bayes 估计即为 Bayes 估计. 下面求解风险函数为常数, 令 $c = \frac{1}{a+b+2n}$, 则

$$\begin{aligned} \mathcal{R}(\theta) &= \mathbb{E} \left[\frac{a+n\bar{X}}{a+b+2n} - \theta \right]^2 = \mathbb{E} [cn\bar{X} + ac - \theta]^2 = \mathbb{E} [c(n\bar{X} - 2n\theta) + (2nc - 1)\theta + ac]^2 \\ &= c^2 \text{Var}(n\bar{X}) + [(2nc - 1)\theta + ac]^2 = c^2 \cdot 2n\theta(1-\theta) + (2nc - 1)^2 \theta^2 + 2ac(2nc - 1)\theta + a^2 c^2 \\ &= [(2nc - 1)^2 - 2nc^2] \theta^2 + [2nc^2 + 2ac(2nc - 1)] \theta + a^2 c^2 \end{aligned}$$

上式与 θ 无关, 得 $\begin{cases} 2nc^2 = (1 - 2nc)^2, \\ 2nc^2 = 2ac(1 - 2nc). \end{cases} \Rightarrow \begin{cases} c = \frac{1}{\sqrt{2n} + 2n}, \\ a = \sqrt{2n}/2 \end{cases}$ 于是 $b = \sqrt{2n}/2$, 所以 θ 的

Minimax 估计为 $\frac{1 + \sqrt{2n}\bar{X}}{2 + 2\sqrt{2n}}$.