

2023 Spring 《数理统计》平时作业 2

@ 做作的 Morpheus

2023 年 5 月 26 日

题目 1. 设 x_1, x_2, \dots, x_n 是来自二点分布 $b(1, \theta)$ 的一个样本, 寻求 θ 与 $g(\theta) = \theta(1 - \theta)$ 的矩估计.

解答. 由于总体 $X \sim b(1, \theta)$, 所以总体均值

$$\mu = \mathbb{E}[x] = \theta$$

的矩估计为 $\hat{\theta} = \bar{x}$. 相应地, 函数 $g(\theta) = \theta(1 - \theta)$ 的矩估计则为

$$\hat{g}(\theta) = g(\hat{\theta}) = \bar{x}(1 - \bar{x}).$$

□

题目 2. 设 x_1, x_2, \dots, x_n 是来自如下离散均匀分布的一个样本:

$$P(X = k) = 1/N, \quad k = 1, 2, \dots, N$$

寻求正整数 N 的矩估计.

解答. 样本 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 所取自总体 X 的均值为

$$\mu = \mathbb{E}[X] = \sum_{k=1}^N kP(X = k) = \frac{N+1}{2} \implies N = 2\mu - 1.$$

用样本均值 \bar{x} 代替总体均值 μ , 即得到 N 的矩估计

$$\hat{N} = 2\hat{\mu} - 1 = 2\bar{x} - 1.$$

□

题目 3. 设 x_1, x_2, \dots, x_n 是来自均匀分布 $U(0, \theta)$ 的一个样本, 寻求 θ 的矩估计, 并讨论其无偏性与相合性.

解答. 总体 $X \sim U(0, \theta)$ 的均值为

$$\mu = \mathbb{E}[x] = \frac{\theta}{2} \implies \theta = 2\mu.$$

用样本均值 \bar{x} 代替总体均值 μ , 即得到 θ 的矩估计

$$\hat{\theta} = 2\hat{\mu} = 2\bar{x}.$$

- 1) **无偏性:** 由于样本均值 \bar{x} 是总体均值 μ 的无偏估计, 所以 $\hat{\theta} = 2\bar{x}$ 也是 $\theta = 2\mu$ 的无偏估计;
- 2) **相合性:** 由辛钦大数定律 (定理 2.1.1) 知样本均值 \bar{x} 是总体均值 μ 的相合估计, 又由定理 2.1.2 可知在连续变换下 $\hat{\theta} = 2\bar{x}$ 也是 $\theta = 2\mu$ 的无偏估计. □

题目 4. 设 x_1, x_2, \dots, x_n 是来自泊松分布 $P(\lambda)$ 的一个样本, 寻找概率 $P(X = 0)$ 的矩估计, 并讨论其相合性.

解答. 总体 $X \sim P(\lambda)$ 的均值为 $\mu = \mathbb{E}[x] = \lambda$, 矩估计为样本均值 $\hat{\mu} = \bar{x}$, 由辛钦大数定律 (定理 2.1.1) 知其为相合估计. 因此概率 $p = P(X = 0) = e^{-\lambda}$ 的矩估计为

$$\hat{p} = e^{-\hat{\lambda}} = e^{-\bar{x}},$$

且由定理 2.1.2 可知其为相合估计. □

题目 5. 设总体 X 服从参数为 λ 的泊松分布, 从中抽取样本 x_1, x_2, \dots, x_n , 求 λ 的最大似然估计.

解答. 对数似然函数

$$\begin{aligned}\ell(\lambda) &= \sum_{i=1}^n \log p(x_i; \lambda) = \sum_{i=1}^n \log \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \\ &= \sum_{i=1}^n (x_i \log \lambda - \lambda - \log x_i!) \\ &= -n\lambda + n\bar{x} \log \lambda - \sum_{i=1}^n \log x_i!\end{aligned}$$

对 λ 求二阶导数得:

$$\begin{aligned}\frac{\partial \ell}{\partial \lambda} &= -n + \frac{n\bar{x}}{\lambda} = 0 \implies \hat{\lambda} = \bar{x} \\ \frac{\partial^2 \ell}{\partial \lambda^2} &= -\frac{n\bar{x}}{\lambda^2}\end{aligned}$$

由于在 $\lambda = \hat{\lambda}$ 处有 $\frac{\partial^2 \ell}{\partial \lambda^2} < 0$, 所以 $\hat{\lambda} = \bar{x}$ 使 $\ell(\lambda)$ 达到最大, 即为 λ 的最大似然估计. \square

题目 6. 设总体 X 的密度函数为:

$$p(x; \beta) = (\beta + 1)x^\beta, \quad 0 < x < 1$$

其中, 未知参数 $\beta > -1$, 从中获得样本 x_1, x_2, \dots, x_n , 求参数 β 的最大似然估计与矩估计, 它们是否相同? 今获得的样本观察值为:

0.30 0.80 0.47 0.35 0.62 0.55

试分别求 β 的两个估计值.

解答.

1) 对数似然函数

$$\begin{aligned}\ell(\beta) &= \sum_{i=1}^n \log p(x_i; \beta) = \sum_{i=1}^n \log(\beta + 1) + \beta \log x_i \\ &= n \log(\beta + 1) + \beta \sum_{i=1}^n \log x_i\end{aligned}$$

对 β 求二阶导数:

$$\begin{aligned}\frac{\partial \ell}{\partial \beta} &= \frac{n}{\beta + 1} + \sum_{i=1}^n \log x_i = 0 \implies \hat{\beta}_1 = -\frac{n}{\sum_{i=1}^n \log x_i} - 1 \\ \frac{\partial^2 \ell}{\partial \beta^2} &= -\frac{n}{(\beta + 1)^2}\end{aligned}$$

由于在 $\beta = \hat{\beta}_1$ 处有 $\frac{\partial^2 \ell}{\partial \beta^2} < 0$, 所以 $\hat{\beta}_1 = -\frac{n}{\sum_{i=1}^n \log x_i} - 1$ 使 $\ell(\beta)$ 达到最大, 即为 β 的最大似然估计.

2) 总体 X 的均值为

$$\begin{aligned}\mu = \mathbb{E}[x] &= \int_0^1 xp(x; \beta) dx = \int_0^1 (\beta + 1)x^{\beta+1} dx \\ &= \frac{\beta + 1}{\beta + 2} x^{\beta+2} \Big|_0^1 = \frac{\beta + 1}{\beta + 2} = 1 - \frac{1}{\beta + 2}\end{aligned}$$

即

$$\beta = \frac{2\mu - 1}{1 - \mu}$$

用样本均值 \bar{x} 代替总体均值 μ , 得到 β 的矩估计

$$\hat{\beta}_2 = \frac{2\bar{x} - 1}{1 - \bar{x}} = \frac{1}{1 - \bar{x}} - 2.$$

3) 要证 $\hat{\beta}_1 \neq \hat{\beta}_2$, 只需取一个反例即可, 题设样本即为一个反例, 另外还可以取

$$x_1 = \cdots = x_n = \frac{1}{e} \implies \hat{\beta}_1 = 0, \hat{\beta}_2 = \frac{e}{e-1} - 2 \neq 0.$$

4) 将 $(x_1, \dots, x_6) = (0.30, 0.80, 0.47, 0.35, 0.62, 0.55)$ 代入计算可得

$$\hat{\beta}_1 = -\frac{6}{\sum_{i=1}^6 \log x_i} - 1 = 0.39, \quad \hat{\beta}_2 = \frac{2 \sum_{i=1}^6 x_i - 6}{6 - \sum_{i=1}^6 x_i} = 0.0619$$

□

题目 7. 设 x_1, x_2, \dots, x_n 与 y_1, y_2, \dots, y_m 分别是来自 $\mathcal{N}(\mu_1, \sigma^2)$ 与 $\mathcal{N}(\mu_2, \sigma^2)$ 的两个独立样本, 试求 μ_1, μ_2, σ^2 的最大似然估计.

解答. 似然函数

$$\begin{aligned} L(\mu_1, \mu_2, \sigma^2) &= \prod_{i=1}^n \mathcal{N}(x_i; \mu_1, \sigma^2) \prod_{j=1}^m \mathcal{N}(y_j; \mu_2, \sigma^2) \\ &= (2\pi\sigma^2)^{-\frac{n+m}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{j=1}^m (y_j - \mu_2)^2 \right] \right\} \end{aligned}$$

对数似然函数

$$\begin{aligned} \ell(\mu_1, \mu_2, \sigma^2) &= -\frac{n+m}{2} \log(2\pi) - \frac{n+m}{2} \log \sigma^2 \\ &\quad - \frac{1}{2\sigma^2} \left[\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{j=1}^m (y_j - \mu_2)^2 \right] \end{aligned}$$

1) 对 μ_1 求偏导数:

$$\frac{\partial \ell}{\partial \mu_1} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu_1) = \frac{n}{\sigma^2} (\bar{x} - \mu_1) = 0 \implies \hat{\mu}_1 = \bar{x}$$

$$\frac{\partial^2 \ell}{\partial \mu_1^2} = -\frac{n}{\sigma^2} < 0$$

由于在 $\mu_1 = \hat{\mu}_1$ 处有 $\frac{\partial^2 \ell}{\partial \mu_1^2} < 0$, 所以在 μ_2, σ^2 确定时 $\hat{\mu}_1 = \bar{x}$ 使 $\ell(\mu_1, \mu_2, \sigma^2)$ 达到最大, 即为 μ_1 的最大似然估计;

2) 对 μ_2 求偏导数:

$$\frac{\partial \ell}{\partial \mu_2} = \frac{1}{\sigma^2} \sum_{j=1}^m (y_j - \mu_2) = \frac{m}{\sigma^2} (\bar{y} - \mu_2) = 0 \implies \hat{\mu}_2 = \bar{y}$$

$$\frac{\partial^2 \ell}{\partial \mu_2^2} = -\frac{m}{\sigma^2} < 0$$

由于在 $\mu_2 = \hat{\mu}_2$ 处有 $\frac{\partial^2 \ell}{\partial \mu_2^2} < 0$, 所以在 μ_1, σ^2 确定时 $\hat{\mu}_2 = \bar{y}$ 使 $\ell(\mu_1, \mu_2, \sigma^2)$ 达到最大, 即为 μ_2 的最大似然估计;

3) 对 σ^2 求偏导数:

$$\begin{aligned} \frac{\partial \ell}{\partial (\sigma^2)} &= -\frac{n+m}{2\sigma^2} + \frac{1}{2\sigma^4} \left[\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{j=1}^m (y_j - \mu_2)^2 \right] = 0 \\ \Rightarrow \hat{\sigma}^2 &= \frac{1}{n+m} \left[\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{j=1}^m (y_j - \mu_2)^2 \right] \end{aligned}$$

求二阶偏导数:

$$\frac{\partial^2 \ell}{\partial (\sigma^2)^2} = \frac{n+m}{2\sigma^4} - \frac{1}{\sigma^6} \left[\sum_{i=1}^n (x_i - \mu_1)^2 + \sum_{j=1}^m (y_j - \mu_2)^2 \right]$$

将 $\hat{\sigma}^2$ 代入可得

$$\left. \frac{\partial^2 \ell}{\partial (\sigma^2)^2} \right|_{\hat{\sigma}^2} = (n+m) \left(\frac{1}{2\hat{\sigma}^2} - \frac{1}{\hat{\sigma}^4} \right) = -\frac{n+m}{2\hat{\sigma}^2} < 0$$

所以在 μ_1, μ_2 确定时 $\hat{\sigma}^2$ 使 $\ell(\mu_1, \mu_2, \sigma^2)$ 达到最大, 即为 σ^2 的最大似然估计. \square

题目 8. 设炮弹着落点 (x, y) 离目标 (原点) 的距离为 $z = \sqrt{x^2 + y^2}$, 若设 x 与 y 为独立同分布的随机变量, 其共同分布为 $\mathcal{N}(0, \sigma^2)$, 可得 z 的分布密度为:

$$p(z) = \frac{z}{\sigma^2} \exp \left\{ -\frac{z^2}{2\sigma^2} \right\}, \quad z > 0$$

这个分布称为瑞利分布.

- 1) 设 z_1, z_2, \dots, z_n 为来自上述瑞利分布的一个样本, 求 σ^2 的 MLE, 证明它是 σ^2 的无偏估计;

2) 求瑞利分布中 σ^2 的费希尔信息量 $I(\sigma^2)$;

3) 给出 MLE $\hat{\sigma}^2$ 的渐近正态分布.

解答.

1) 似然函数

$$L(\sigma^2) = \prod_{i=1}^n p(z_i; \sigma^2) = \sigma^{-2n} \prod_{i=1}^n z_i \exp \left\{ -\frac{z_i^2}{2\sigma^2} \right\}$$

对数似然函数

$$\ell(\sigma^2) = -n \log \sigma^2 + \sum_{i=1}^n \log z_i - \frac{1}{2\sigma^2} \sum_{i=1}^n z_i^2$$

求二阶偏导数:

$$\frac{\partial \ell}{\partial (\sigma^2)} = -\frac{n}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n z_i^2 = 0 \implies \hat{\sigma}^2 = \frac{1}{2n} \sum_{i=1}^n z_i^2$$

$$\frac{\partial^2 \ell}{\partial (\sigma^2)^2} = \frac{n}{\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^n z_i^2$$

代入 $\hat{\sigma}^2$ 可得

$$\left. \frac{\partial^2 \ell}{\partial (\sigma^2)^2} \right|_{\hat{\sigma}^2} = -\frac{n}{\hat{\sigma}^4} < 0$$

所以 $\hat{\sigma}^2$ 为 σ^2 的最大似然估计, 且满足

$$\mathbb{E}[\hat{\sigma}^2] = \frac{1}{2n} \sum_{i=1}^n \mathbb{E}[z_i^2] = \frac{1}{2n} \sum_{i=1}^n (\mathbb{E}[x_i^2] + \mathbb{E}[y_i^2]) = \frac{1}{2n} \cdot 2n\sigma^2 = \sigma^2,$$

即为 σ^2 的无偏估计.

- 2) 可以验证瑞利分布属于 Cramer-Rao 正则族 (此处省略), 而费希尔信息量 $I(\sigma^2)$ 的计算如下:

$$\begin{aligned}\log p(z; \sigma^2) &= -\log \sigma^2 + \log z - \frac{z^2}{2\sigma^2} \\ \implies \frac{\partial \log p}{\partial \sigma^2} &= -\frac{1}{\sigma^2} + \frac{z^2}{2\sigma^4} \\ \implies \frac{\partial^2 \log p}{\partial (\sigma^2)^2} &= \frac{1}{\sigma^4} - \frac{z^2}{\sigma^6}\end{aligned}$$

所以

$$I(\sigma^2) = -\mathbb{E}_z \left[\frac{\partial^2 \log p}{\partial (\sigma^2)^2} \right] = \mathbb{E}_z \left[-\frac{1}{\sigma^4} + \frac{z^2}{\sigma^6} \right] = -\frac{1}{\sigma^4} + \frac{2\sigma^2}{\sigma^6} = \frac{1}{\sigma^4}$$

- 3) 由定理 2.2.2 可知, $\hat{\sigma}^2 \sim AN\left(\sigma^2, \frac{\sigma^4}{n}\right)$. □

题目 8 的注记. 上述费希尔信息量的计算利用了如下结论: 设 $p(x; \theta)$ 为 Cramer-Rao 正则族分布, 若其二阶偏导数 $\frac{\partial^2 \log p}{\partial \theta^2}$ 对一切 θ 存在, 则其费希尔信息量为

$$I(\theta) = -\mathbb{E}_x \left[\frac{\partial^2 \log p}{\partial \theta^2} \right]$$

事实上, 该结论往往可以有效简化费希尔信息量的计算 (在以下题目中也会使用), 其证明也并不困难, 首先需要注意到

$$\begin{aligned}\mathbb{E}_x[S_\theta] &= \mathbb{E}_x \left[\frac{\partial \log p}{\partial \theta} \right] = \int_{\mathbb{R}} \frac{\partial p}{\partial \theta} \cdot \frac{1}{p} \cdot p \, dx \\ &= \int_{\mathbb{R}} \frac{\partial p}{\partial \theta} \, dx = \frac{\partial}{\partial \theta} \int_{\mathbb{R}} p \, dx = \frac{\partial}{\partial \theta} 1 = 0.\end{aligned}$$

其中我们记 $S_\theta = \frac{\partial \log p}{\partial \theta}$, 于是有

$$\begin{aligned} 0 &= \frac{\partial \mathbb{E}_x[S_\theta]}{\partial \theta} = \int_{\mathbb{R}} \frac{\partial(S_\theta p)}{\partial \theta} dx \\ &= \int_{\mathbb{R}} \frac{\partial S_\theta}{\partial \theta} p dx + \int_{\mathbb{R}} S_\theta \frac{\partial p}{\partial \theta} dx \\ &= \int_{\mathbb{R}} \frac{\partial^2 \log p}{\partial \theta^2} p dx + \int_{\mathbb{R}} S_\theta \frac{\partial \log p}{\partial \theta} p dx \\ &= \mathbb{E}_x \left[\frac{\partial^2 \log p}{\partial \theta^2} \right] + \mathbb{E}_x[S_\theta^2] \end{aligned}$$

其中 $\mathbb{E}_x[S_\theta^2] = I(\theta)$, 因此

$$I(\theta) = -\mathbb{E}_x \left[\frac{\partial^2 \log p}{\partial \theta^2} \right].$$

题目 9. 设某种电器的寿命 (单位: 小时) 服从指数分布 $\exp(1/\theta)$. 现有 10 件此种电器同时参加寿命试验, 已知 2 件在 110 和 170 小时先后发生失效, 其余的在 200 小时停止试验前再没有发生失效, 试求其平均失效时间的最大似然估计.

解答. 记该种电器的寿命为 $X \sim \exp(1/\theta)$, 其样本 x_1, \dots, x_{10} 满足

$$p(x_i; \theta) = \frac{1}{\theta} \exp \left\{ -\frac{x_i}{\theta} \right\}, \quad x_i \geq 0, \quad i = 1, \dots, 10$$

令 $n = 10$, 则对数似然函数

$$\ell(\theta) = \sum_{i=1}^n \log p(x_i; \theta) = \sum_{i=1}^n -\log \theta - \frac{x_i}{\theta} = -n \log \theta - \frac{n}{\theta} \bar{x}$$

对 θ 求二阶偏导数:

$$\begin{aligned} \frac{\partial \ell}{\partial \theta} &= -\frac{n}{\theta} + \frac{n\bar{x}}{\theta^2} = 0 \implies \hat{\theta} = \bar{x} \\ \frac{\partial^2 \ell}{\partial \theta^2} &= \frac{n}{\theta^2} - \frac{2n\bar{x}}{\theta^3} \implies \left. \frac{\partial^2 \ell}{\partial \theta^2} \right|_{\hat{\theta}} = -\frac{n}{\hat{\theta}^2} < 0 \end{aligned}$$

由于 $\hat{\theta} = \bar{x}$ 使 $\ell(\theta)$ 达到最大, 即为 θ 的最大似然估计. 代入 $n = 10$ 及样本 $x_1 = 110, x_2 = 170, x_3 = \cdots = x_{10} = 200$, 可得最大似然估计

$$\hat{\theta} = \bar{x} = \frac{1}{10}(110 + 170 + 8 \times 200) = 188 \text{ (h)}$$

□

题目 10. 设 $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 是参数 θ 的两个无偏估计, 且相互独立, 其方差分别为 $\text{Var}(\hat{\theta}_1) = \sigma_1^2, \text{Var}(\hat{\theta}_2) = \sigma_2^2$. 要求:

- 1) 对任意 $\alpha (0 < \alpha < 1)$, 证明 $\hat{\theta}_\alpha = \alpha\hat{\theta}_1 + (1 - \alpha)\hat{\theta}_2$ 是 θ 的无偏估计;
- 2) α 为何值时, 可得 $\hat{\theta}_\alpha$ 的方差最小?

解答.

- 1) 对任意的 $0 < \alpha < 1$, 有

$$\mathbb{E}[\hat{\theta}_\alpha] = \alpha\mathbb{E}[\hat{\theta}_1] + (1 - \alpha)\mathbb{E}[\hat{\theta}_2] = \alpha\theta + (1 - \alpha)\theta = \theta,$$

即 $\hat{\theta}_\alpha$ 是 θ 的无偏估计.

- 2) 由于 $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 相互独立,

$$\begin{aligned} \text{Var}[\hat{\theta}_\alpha] &= \alpha^2 \text{Var}[\hat{\theta}_1] + (1 - \alpha)^2 \text{Var}[\hat{\theta}_2] \\ &= \alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2 \\ &= (\sigma_1^2 + \sigma_2^2)\alpha^2 - 2\sigma_2^2\alpha + \sigma_2^2 \\ &= (\sigma_1^2 + \sigma_2^2) \left(\alpha - \frac{\alpha_2^2}{\sigma_1^2 + \sigma_2^2} \right)^2 + \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \end{aligned}$$

于是由二次函数的性质可知, 当 $\alpha = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ 时, $\hat{\theta}_\alpha$ 的方差最小. □

题目 11. 设 x_1, x_2, \dots, x_n 是来自均匀分布 $U(\theta, \theta + 1)$ 的一个样本. 要求:

- 1) 验证 $\hat{\theta}_1 = \bar{x} - \frac{1}{2}$, $\hat{\theta}_2 = x_{(1)} - \frac{1}{n+1}$, $\hat{\theta}_3 = x_{(n)} - \frac{n}{n+1}$ 都是 θ 的无偏估计;
- 2) 比较这三个估计的有效性.

解答.

- 1) 分别计算 $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ 的均值:

$$1. \mathbb{E}[\hat{\theta}_1] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[x_i] - \frac{1}{2} = \theta + \frac{1}{2} - \frac{1}{2} = \theta$$

2. 次序统计量 $x_{(1)}$ 的密度函数为

$$p_1(x) = n[1 - F(x)]^{n-1}p(x) = n(1 - x + \theta)^{n-1}$$

其中

$$p(x) = 1, F(x) = x - \theta, \quad x \in [\theta, \theta + 1]$$

所以 $x_{(1)}$ 的均值为

$$\begin{aligned} \mathbb{E}[x_{(1)}] &= \int_{\theta}^{\theta+1} nx(\theta + 1 - x)^{n-1} dx \\ &= \int_0^1 n(\theta + 1 - x)x^{n-1} dx \\ &= n(\theta + 1) \int_0^1 x^{n-1} dx - n \int_0^1 x^n dx \\ &= \theta + 1 - \frac{n}{n+1} = \theta + \frac{1}{n+1} \end{aligned}$$

因此, $\hat{\theta}_2$ 的均值为

$$\mathbb{E}[\hat{\theta}_2] = \mathbb{E}[x_{(1)}] - \frac{1}{n+1} = \theta$$

3. 次序统计量 $x_{(n)}$ 的密度函数为

$$p_n(x) = n[F(x)]^{n-1}p(x) = n(x - \theta)^{n-1}$$

其均值为

$$\begin{aligned}
 \mathbb{E}[x_{(n)}] &= \int_{\theta}^{\theta+1} nx(x-\theta)^{n-1} dx \\
 &= \int_0^1 n(x+\theta)x^{n-1} dx \\
 &= n \int_0^1 x^n dx + n\theta \int_0^1 x^{n-1} dx \\
 &= \theta + \frac{n}{n+1}
 \end{aligned}$$

因此, $\hat{\theta}_3$ 的均值为

$$\mathbb{E}[\hat{\theta}_3] = \mathbb{E}[x_{(n)}] - \frac{n}{n+1} = \theta$$

由于 $\mathbb{E}[\hat{\theta}_1] = \mathbb{E}[\hat{\theta}_2] = \mathbb{E}[\hat{\theta}_3] = \theta$, 故 $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ 均为 θ 的无偏估计.

2) 分别计算 $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ 的方差:

$$1. \text{Var}[\hat{\theta}_1] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}[x_i] = \frac{1}{12n}$$

2. 由于 $\text{Var}[\hat{\theta}_2] = \text{Var}[x_{(1)}]$, 所以

$$\begin{aligned}
 \text{Var}[\hat{\theta}_2] &= \int_{\theta}^{\theta+1} n \left(x - \theta - \frac{1}{n+1} \right)^2 (\theta+1-x)^{n-1} dx \\
 &= n \int_0^1 \left(\frac{n}{n+1} - x \right)^2 x^{n-1} dx \\
 &= n \left[\frac{n^2}{(n+1)^2} \int_0^1 x^{n-1} dx - \frac{2n}{n+1} \int_0^1 x^n dx + \int_0^1 x^{n+1} dx \right] \\
 &= n \left[\frac{n}{(n+1)^2} - \frac{2n}{(n+1)^2} + \frac{1}{n+2} \right] \\
 &= n \cdot \frac{(n+1)^2 - n(n+2)}{(n+1)^2(n+2)} = \frac{n}{(n+1)^2(n+2)}
 \end{aligned}$$

3. 由于 $\text{Var}[\hat{\theta}_3] = \text{Var}[x_{(n)}]$, 所以

$$\begin{aligned}
 \text{Var}[\hat{\theta}_3] &= \int_{\theta}^{\theta+1} n \left(x - \theta - \frac{n}{n+1} \right)^2 (x-\theta)^{n-1} dx \\
 &= n \int_0^1 \left(x - \frac{n}{n+1} \right)^2 x^{n-1} dx = \frac{n}{(n+1)^2(n+2)}
 \end{aligned}$$

其中 $\text{Var}[\hat{\theta}_2] = \text{Var}[\hat{\theta}_3]$. 另外, 我们通过作差:

$$\begin{aligned}\text{Var}[\hat{\theta}_1] - \text{Var}[\hat{\theta}_2] &= \frac{1}{12n} - \frac{n}{(n+1)^2(n+2)} \\ &= \frac{1}{12n(n+1)^2(n+2)} [(n+1)^2(n+2) - 12n^2] \\ &= \frac{(n-1)(n^2-7n-2)}{12n(n+1)^2(n+2)} \begin{cases} < 0, & n \leq 7 \\ > 0, & n \geq 8 \end{cases}\end{aligned}$$

可以发现: 当 $n \leq 7$ 时, $\text{Var}[\hat{\theta}_1] < \text{Var}[\hat{\theta}_2] = \text{Var}[\hat{\theta}_3]$, 即 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效, 而 $\hat{\theta}_2$ 和 $\hat{\theta}_3$ 同等有效; 当 $n \geq 8$ 时, $\text{Var}[\hat{\theta}_1] > \text{Var}[\hat{\theta}_2] = \text{Var}[\hat{\theta}_3]$, 即 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 有效, 而 $\hat{\theta}_2$ 和 $\hat{\theta}_3$ 同等有效. \square

题目 12. 设 x_1, x_2, \dots, x_n 是来自指数分布 $\exp(1/\theta)$ 的一个样本, 试证 $\hat{\theta}_1 = \bar{x}$ 与 $\hat{\theta}_2 = nx_{(1)}$ 都是 θ 的无偏估计, 并比较其有效性.

解答.

1) 分别计算 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 的均值:

1. $\hat{\theta}_1$ 的均值为: $\mathbb{E}[\hat{\theta}_1] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[x_i] = \theta$

2. 次序统计量 $x_{(1)}$ 的密度函数为

$$p_1(x) = n[1 - F(x)]^{n-1}p(x) = \frac{n}{\theta} \exp\left\{-\frac{nx}{\theta}\right\}$$

其中

$$p(x) = \frac{1}{\theta} \exp\left\{-\frac{x}{\theta}\right\}, \quad F(x) = 1 - \exp\left\{-\frac{x}{\theta}\right\}, \quad x \geq 0$$

所以 $x_{(1)}$ 的均值为

$$\begin{aligned}\mathbb{E}[x_{(1)}] &= \int_0^{+\infty} x \frac{n}{\theta} e^{-\frac{nx}{\theta}} dx = \frac{\theta}{n} \int_0^{+\infty} x e^{-x} dx \\ &= \frac{\theta}{n} \int_0^{+\infty} (-x) de^{-x} = \frac{\theta}{n} \left(-xe^{-x} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-x} dx \right) = \frac{\theta}{n}\end{aligned}$$

因此, $\hat{\theta}_2$ 的均值为:

$$\mathbb{E}[\hat{\theta}_2] = n\mathbb{E}[x_{(1)}] = \theta$$

由于 $\mathbb{E}[\hat{\theta}_1] = \mathbb{E}[\hat{\theta}_2] = \theta$, 所以 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 均为 θ 的无偏估计.

2) 分别计算 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 的方差:

$$1. \text{Var}[\hat{\theta}_1] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}[x_i] = \frac{1}{n^2} \cdot n\theta^2 = \frac{\theta^2}{n}$$

2. 先计算次序统计量 $x_{(1)}$ 的方差:

$$\begin{aligned} \text{Var}[x_{(1)}] &= \int_0^{+\infty} \left(x - \frac{\theta}{n}\right)^2 \frac{n}{\theta} e^{-\frac{nx}{\theta}} dx \\ &= \frac{\theta^2}{n^2} \int_0^{+\infty} (x-1)^2 e^{-x} dx \\ &= \frac{\theta^2}{n^2} \int_0^{+\infty} -(x-1)^2 de^{-x} \\ &= \frac{\theta^2}{n^2} \left[-(x-1)^2 e^{-x} \Big|_0^{+\infty} + \int_0^{+\infty} 2(x-1)e^{-x} dx \right] \\ &= \frac{\theta^2}{n^2} (1 + 2 \times 1 - 2 \times 1) = \frac{\theta^2}{n^2} \end{aligned}$$

于是 $\hat{\theta}_2$ 的方差为:

$$\text{Var}[\hat{\theta}_2] = n^2 \text{Var}[x_{(1)}] = \theta^2$$

当 $n=1$ 时, $\text{Var}[\hat{\theta}_1] = \text{Var}[\hat{\theta}_2] = \theta^2$; 当 $n \geq 2$ 时, $\text{Var}[\hat{\theta}_1] < \text{Var}[\hat{\theta}_2]$.

因此, $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效. \square

题目 13. 设 x_1, x_2, \dots, x_n 是来自指数分布 $\exp(1/\theta)$ 的一个样本, 求 c 使 $c\bar{x}$ 在均方误差准则下是 θ 的最优估计.

解答. 由于 $\mathbb{E}[x_i] = \theta$, $\text{Var}[x_i] = \theta^2$, 所以 $c\bar{x}$ 的均方误差为

$$\begin{aligned} MSE(c\bar{x}) &= \mathbb{E}[(c\bar{x} - \theta)^2] = \text{Var}[c\bar{x}] + (\mathbb{E}[c\bar{x}] - \theta)^2 \\ &= \frac{c^2}{n^2} \sum_{i=1}^n \text{Var}[x_i] + \left(\frac{c}{n} \sum_{i=1}^n \mathbb{E}[x_i] - \theta \right)^2 \\ &= \frac{c^2}{n} \theta^2 + \theta^2 (c - 1)^2 \\ &= \frac{n+1}{n} \theta^2 \left(c - \frac{n}{n+1} \right)^2 + \frac{\theta^2}{n+1} \end{aligned}$$

因此当 $c = \frac{n}{n+1}$ 时, $c\bar{x}$ 在均方误差准则下是 θ 的最优估计. \square

题目 14. 检验下列分布族的完备性:

- 1) 泊松分布族;
- 2) 几何分布族;
- 3) 均匀分布族 $\{U(0, \theta) : \theta > 0\}$;
- 4) 伽玛分布族 $\{Ga(\alpha, \lambda) : \alpha > 0, \lambda > 0\}$.

解答.

- 1) 设 $X \sim P(\lambda)$, 其中 $\lambda > 0$, 则

$$P_\lambda(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (x = 0, 1, 2, \dots)$$

假设对任意 $\lambda > 0$ 有 $\mathbb{E}_\lambda[\varphi(x)] = 0$, 即

$$\sum_{x=0}^{\infty} \varphi(x) \cdot \frac{\lambda^x e^{-\lambda}}{x!} = 0$$

由于 $\frac{\lambda^x e^{-\lambda}}{x!} > 0$, 于是有

$$\varphi(x) = 0, \quad x = 0, 1, 2, \dots \implies P_\lambda \{\varphi(x) = 0\} = 1$$

因此泊松分布族是完备的;

2) 设 $X \sim Ge(p)$, 其中 $0 < p \leq 1$, 则

$$P(X = x) = (1 - p)^{x-1}p \quad (x = 1, 2, \dots)$$

假设对任意 $0 < p \leq 1$ 有 $\mathbb{E}_p[\varphi(x)] = 0$, 即

$$\sum_{x=0}^{\infty} \varphi(x)(1 - p)^{x-1}p = 0$$

由于 $(1 - p)^{x-1}p > 0$ 当 $0 < p < 1$ 时成立, 于是有

$$\varphi(x) = 0, \quad x = 1, 2, \dots \implies P_p \{\varphi(x) = 0\} = 1$$

因此几何分布族是完备的;

3) 设 $X \sim U(0, \theta)$, 其中 $\theta > 0$, 则

$$p_\theta(x) = \frac{1}{\theta} \quad (0 < x < \theta)$$

假设对任意 $\theta > 0$ 有 $\mathbb{E}_\theta[\varphi(x)] = 0$, 于是由微积分基本定理:

$$\int_0^\theta \varphi(x) dx = 0 \implies P_\theta \{\varphi(x) = 0\} = 1$$

因此均匀分布族是完备的;

4) 设 $X \sim Ga(\alpha, \lambda)$, 其中 $\alpha > 0, \lambda > 0$, 则

$$p_{\alpha, \lambda}(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad (x \geq 0)$$

假设对任意 $\alpha > 0, \lambda > 0$ 有 $\mathbb{E}_{\alpha, \lambda}[\varphi(x)] = 0$, 即

$$\int_0^{+\infty} \varphi(x) \cdot \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx = 0$$

消去非零常数, 则有

$$\int_0^{+\infty} \varphi(x) x^{\alpha-1} e^{-\lambda x} dx = \mathcal{L} \{\varphi(x) x^{\alpha-1}\}(\lambda) = 0.$$

由 Laplace 变换的唯一性可知

$$P_\lambda \{\varphi(x) x^{\alpha-1} = 0\} = 1 \implies P_{\alpha, \lambda} \{\varphi(x) = 0\} = 1$$

因此伽玛分布族是完备的. □

题目 15. 设 T 是 $g(\theta)$ 的 UMVUE, \hat{g} 是 $g(\theta)$ 的无偏估计, 证明: 若 $\text{Var}(\hat{g}) < +\infty$, 则 $\text{Cov}(T, \hat{g}) \geq 0$.

解答. 记 $U = T - \hat{g}$, 则

$$\mathbb{E}_\theta[U] = \mathbb{E}_\theta[T] - \mathbb{E}_\theta[\hat{g}] = g(\theta) - g(\theta) = 0 \implies U \in \mathcal{U}_0$$

其中 \mathcal{U}_0 表示零的无偏估计类. 由于 $\text{Var}_\theta[U] = \text{Var}_\theta[T] + \text{Var}_\theta[\hat{g}] < +\infty$, 于是由定理 2.3.1 可知

$$\text{Cov}_\theta(T, U) = \text{Cov}_\theta(T, T - \hat{g}) = \text{Var}_\theta[T] - \text{Cov}_\theta(T, \hat{g}) = 0$$

因此 $\text{Cov}_\theta(T, \hat{g}) = \text{Var}_\theta[T] \geq 0$. □

题目 16. 设 x_1, x_2, \dots, x_n 是来自如下密度函数的一个样本

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

- 1) 求 $g(\theta) = 1/\theta$ 的最大似然估计 $\hat{g}(\mathbf{x})$;
- 2) 验证 $\hat{g}(\mathbf{x})$ 是 $g(\theta)$ 的无偏估计;
- 3) 求该分布的费希尔信息量 $I(\theta)$;
- 4) 考察 $\hat{g}(\mathbf{x})$ 的方差是否达到 C-R 下界.

解答.

- 1) 对数似然函数

$$\ell(\theta) = \sum_{i=1}^n \log \theta + (\theta - 1) \log x_i = n \log \theta + (\theta - 1) \sum_{i=1}^n \log x_i$$

对 θ 求二阶偏导数:

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = 0 \implies \hat{\theta} = -\frac{n}{\sum_{i=1}^n \log x_i}$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{n}{\theta^2} < 0$$

所以 $\hat{\theta}$ 为 θ 的最大似然估计, 于是 $g(\theta) = 1/\theta$ 的最大似然估计为

$$\hat{g}(\mathbf{x}) = g(\hat{\theta}) = \frac{1}{\hat{\theta}} = -\frac{1}{n} \sum_{i=1}^n \log x_i$$

2) 对于上述总体 X , 先计算 $\log x$ 的期望:

$$\mathbb{E}_x[\log x] = \int_0^1 \log x \, dx^\theta = x^\theta \log x \Big|_0^1 - \int_0^1 x^{\theta-1} \, dx = -\frac{1}{\theta}$$

于是 $\hat{g}(\mathbf{x})$ 的期望为

$$\mathbb{E}_x[\hat{g}(\mathbf{x})] = -\frac{1}{n} \sum_{i=1}^n \mathbb{E}_{x_i}[\log x_i] = \frac{1}{\theta} = g(\theta)$$

即 $\hat{g}(\mathbf{x})$ 是 $g(\theta)$ 的无偏估计.

3) 由于

$$\begin{aligned} \log p(x; \theta) &= \log \theta + (\theta - 1) \log x \\ \implies \frac{\partial \log p}{\partial \theta} &= \frac{1}{\theta} + \log x \\ \implies \frac{\partial^2 \log p}{\partial \theta^2} &= -\frac{1}{\theta^2} \end{aligned}$$

所以费希尔信息量为

$$I(\theta) = -\mathbb{E}_x \left[\frac{\partial^2 \log p}{\partial \theta^2} \right] = \frac{1}{\theta^2}$$

4) 先计算 $\log x$ 的方差:

$$\begin{aligned} \text{Var}_x[\log x] &= \mathbb{E}_x \left[\left(\log x + \frac{1}{\theta} \right)^2 \right] \\ &= \mathbb{E}_x \left[\left(\frac{\partial \log p}{\partial \theta} \right)^2 \right] = I(\theta) = \frac{1}{\theta^2} \end{aligned}$$

于是 $\hat{g}(\mathbf{x})$ 的方差为

$$\text{Var}_{\mathbf{x}}[\hat{g}(\mathbf{x})] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}_{x_i}[\log x_i] = \frac{1}{n^2} \cdot \frac{n}{\theta^2} = \frac{1}{n\theta^2}$$

另外, C-R 下界为

$$\frac{[g'(\theta)]^2}{nI(\theta)} = \frac{1/\theta^4}{n/\theta^2} = \frac{1}{n\theta^2}$$

因此, $\hat{g}(\mathbf{x})$ 的方差达到了 C-R 下界. □

题目 16 的注记. 注意到总体 X 的密度函数, 我们可以直接令 $Y = -\log X \Leftrightarrow X = \exp(-Y)$, 其密度函数为

$$p(y; \theta) = \theta(e^{-y})^{\theta-1} \cdot |-e^{-y}| = \theta e^{-\theta y}, \quad y > 0$$

即 $Y = -\log X \sim \text{Exp}(\theta)$, 由此我们可以直接得到

$$\mathbb{E}_{\mathbf{x}}[-\log X] = \frac{1}{\theta}, \quad \text{Var}_{\mathbf{x}}[-\log X] = \frac{1}{\theta^2}$$

题目 17. 设 x_1, x_2, \dots, x_n 是来自正态总体 $\mathcal{N}(\theta, 1)$ 的一个样本, 求 θ^2 的 UMVUE, 并指出它不是 θ^2 的有效估计.

解答.

- 1) 由正态分布的性质, \bar{x} 是 θ 的充分统计量, 且 $\bar{x} \sim \mathcal{N}(\theta, 1/n)$. 因正态分布属于指数分布族, 故 \bar{x} 是 θ 的完备充分统计量, 且满足

$$\mathbb{E}_{\mathbf{x}}[\bar{x}^2] = \text{Var}_{\mathbf{x}}[\bar{x}] + \mathbb{E}_{\mathbf{x}}[\bar{x}]^2 = \frac{1}{n} + \theta^2 \implies \mathbb{E}_{\mathbf{x}}\left[\bar{x}^2 - \frac{1}{n}\right] = \theta^2$$

由定理 2.3.3 可知, $T(\mathbf{x}) = \bar{x}^2 - \frac{1}{n}$ 是 θ^2 的 UMVUE.

- 2) 由于 $\bar{x} \sim \mathcal{N}(\theta, 1/n)$, 所以 $\bar{x} = \theta + \frac{1}{\sqrt{n}}z$, 其中 $z \sim \mathcal{N}(0, 1)$. 下面先计

算 \bar{x}^2 的方差:

$$\begin{aligned}
 \text{Var}_{\mathbf{x}}[\bar{x}^2] &= \mathbb{E}_{\mathbf{x}}[\bar{x}^4] - \mathbb{E}_{\mathbf{x}}[\bar{x}^2]^2 \\
 &= \mathbb{E}_z \left[\frac{z^4}{n^2} + \frac{4z^3}{n^2\sqrt{n}}\theta + \frac{6z^2}{n}\theta^2 + \frac{4z}{\sqrt{n}}\theta^3 + \theta^4 \right] \\
 &\quad - \mathbb{E}_z \left[\frac{z^2}{n} + \frac{2z}{\sqrt{n}}\theta + \theta^2 \right]^2 \\
 &= \left(\frac{3}{n^2} + \frac{6}{n}\theta^2 + \theta^4 \right) - \left(\frac{1}{n} + \theta^2 \right)^2 \\
 &= \frac{4}{n}\theta^2 + \frac{2}{n^2} = \frac{4\theta^2 + 2/n}{n}
 \end{aligned}$$

所以 $T(\mathbf{x})$ 的方差为

$$\text{Var}_{\mathbf{x}}[T(\mathbf{x})] = \text{Var}_{\mathbf{x}}[\bar{x}^2] = \frac{4\theta^2 + 2/n}{n}$$

为求 C-R 下界, 我们先计算费希尔信息量:

$$\begin{aligned}
 \log p(x; \theta) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2}(x - \theta)^2 \\
 \implies \frac{\partial \log p}{\partial \theta} &= x - \theta, \quad \frac{\partial^2 \log p}{\partial \theta^2} = -1 \\
 \implies I(\theta) &= -\mathbb{E}_x \left[\frac{\partial^2 \log p}{\partial \theta^2} \right] = 1
 \end{aligned}$$

令 $g(\theta) = \theta^2$, 则得到 C-R 下界

$$\frac{[g'(\theta)]^2}{nI(\theta)} = \frac{4\theta^2}{n}$$

由于 $\text{Var}_{\mathbf{x}}[T(\mathbf{x})] \neq \frac{4\theta^2}{n}$, 故 $T(\mathbf{x})$ 不是 θ^2 的有效估计. □

题目 17 的注记. 对任意的 $z \sim \mathcal{N}(0, 1)$, 有如下结论:

$$\mathbb{E}_z[z^n] = \begin{cases} 0, & n \text{ odd} \\ (n-1)!!, & n \text{ even} \end{cases}$$

其中 n 为奇数的情形由奇函数的性质直接得出, 而 n 为偶数的情形可以通过递推得到:

$$\begin{aligned}
 \mathbb{E}_z[z^n] &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n \exp\left(-\frac{x^2}{2}\right) dx \\
 &= \frac{1}{\sqrt{2\pi}} \left[-\int_{\mathbb{R}} x^{n-1} d \exp\left(-\frac{x^2}{2}\right) \right] \\
 &= \frac{1}{\sqrt{2\pi}} \left[-x^{n-1} \exp\left(-\frac{x^2}{2}\right) \Big|_{-\infty}^{+\infty} + \int_{\mathbb{R}} (n-1)x^{n-2} \exp\left(-\frac{x^2}{2}\right) dx \right] \\
 &= \frac{1}{\sqrt{2\pi}} \cdot (n-1) \int_{\mathbb{R}} x^{n-2} \exp\left(-\frac{x^2}{2}\right) dx \\
 &= \cdots = \frac{1}{\sqrt{2\pi}} \cdot (n-1)!! \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2}\right) dx = (n-1)!!
 \end{aligned}$$

题目 18. 设随机变量 X 的密度函数为:

$$p(x|\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta < 1$$

从中获得容量为 1 的样本, 观察值记为 x .

- 1) 假如 θ 的先验分布为 $U(0, 1)$, 求 θ 的后验分布;
- 2) 假如 θ 的先验分布为 $\pi(\theta) = 3\theta^2$ ($0 < \theta < 1$), 求 θ 的后验分布.

解答.

- 1) θ 的先验分布:

$$\pi(\theta) = 1, \quad 0 < \theta < 1$$

样本 x 与 θ 的联合分布:

$$h(x, \theta) = \pi(\theta)p(x|\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta < 1$$

样本 x 的边缘分布:

$$m(x) = \int_x^1 \frac{2x}{\theta^2} d\theta = -\frac{2x}{\theta} \Big|_{\theta=x}^{\theta=1} = 2(1-x), \quad 0 < x < 1$$

因此, θ 的后验分布为

$$\pi(\theta|x) = \frac{h(x, \theta)}{m(x)} = \frac{x}{(1-x)\theta^2}, \quad 0 < x < \theta < 1$$

2) θ 的先验分布:

$$\pi(\theta) = 3\theta^2, \quad 0 < \theta < 1$$

样本 x 与 θ 的联合分布:

$$h(x, \theta) = \pi(\theta)p(x|\theta) = 6x, \quad 0 < x < \theta < 1$$

样本 x 的边际分布:

$$m(x) = \int_x^1 6x \, d\theta = 6x\theta \Big|_{\theta=x}^{\theta=1} = 6x(1-x), \quad 0 < x < 1$$

因此, θ 的后验分布为

$$\pi(\theta|x) = \frac{h(x, \theta)}{m(x)} = \frac{1}{1-x}, \quad 0 < x < \theta < 1$$

□

题目 19. 某人每天早上在汽车站等候公共汽车的时间 (单位: 分钟) 服从均匀分布 $U(0, \theta)$, 其中 θ 未知, 设 θ 的先验分布的密度函数为:

$$\pi(\theta) = \frac{192}{\theta^4}, \quad \theta \geq 4$$

假如此人三个早上的等车时间分别为 5, 3, 8 分钟, 求 θ 的后验分布.

解答. 记等待公共汽车的时间为 $X \sim U(0, \theta)$, 从中获得容量为 3 的样本 $\mathbf{x} = (x_1, x_2, x_3)$, 其中 $x_1 = 5, x_2 = 3, x_3 = 8$. 此时, 样本 \mathbf{x} 的联合条件密度函数为:

$$p(\mathbf{x}|\theta) = \prod_{i=1}^3 p(x_i|\theta) = \frac{1}{\theta^3} I_{[8, +\infty)}(\theta), \quad 0 < x_i < \theta$$

样本 \mathbf{x} 与 θ 的联合分布:

$$h(\mathbf{x}, \theta) = \pi(\theta)p(\mathbf{x}|\theta) = \frac{192}{\theta^7} I_{[8, +\infty)}(\theta), \quad 0 < x_i < \theta$$

样本 \mathbf{x} 的边际分布:

$$m(\mathbf{x}) = \int_8^{+\infty} h(\mathbf{x}, \theta) d\theta = -\frac{32}{\theta^6} \Big|_8^{+\infty} = \frac{32}{8^6} = \frac{1}{8192}, \quad x_i > 0$$

因此, θ 的后验分布为

$$\pi(\theta|\mathbf{x}) = \frac{h(\mathbf{x}, \theta)}{m(\mathbf{x})} = \frac{1572864}{\theta^7} I_{[8, +\infty)}(\theta)$$

□

题目 20. 设随机变量 X 服从几何分布, 即

$$P(X = k|\theta) = \theta(1 - \theta)^k, \quad k = 0, 1, 2, \dots$$

其中, 参数 θ 的先验分布为均匀分布 $U(0, 1)$.

- 1) 若只对 X 作一次观察, 观察值为 3, 求 θ 的贝叶斯估计;
- 2) 若对 X 作三次观察, 观察值为 2, 3, 5, 求 θ 的贝叶斯估计.

解答. 设样本 $\mathbf{x} = (x_1, x_2, \dots, x_n)$, 其联合条件密度函数为:

$$P(\mathbf{X} = \mathbf{x}|\theta) = \prod_{i=1}^n \theta(1 - \theta)^{x_i} = \theta^n (1 - \theta)^{\sum_{i=1}^n x_i}, \quad x_i = 0, 1, 2, \dots$$

样本 \mathbf{x} 与 θ 的联合分布:

$$h(\mathbf{x}, \theta) = \pi(\theta)p(\mathbf{x}|\theta) = \theta^n (1 - \theta)^{\sum_{i=1}^n x_i}, \quad x_i = 0, 1, 2, \dots, \quad 0 < \theta < 1$$

样本 \mathbf{x} 的边际分布:

$$m(\mathbf{x}) = \int_0^1 h(\mathbf{x}, \theta) d\theta = B\left(n + 1, \sum_{i=1}^n x_i + 1\right), \quad x_i = 0, 1, 2, \dots$$

于是, θ 的后验分布为

$$\pi(\theta|\mathbf{x}) = \frac{h(\mathbf{x}, \theta)}{m(\mathbf{x})} = \frac{\theta^n(1-\theta)^{\sum_{i=1}^n x_i}}{B(n+1, \sum_{i=1}^n x_i + 1)} = Be\left(n+1, \sum_{i=1}^n x_i + 1\right)$$

由此得到 θ 的贝叶斯估计

$$\hat{\theta}_B = \mathbb{E}_\theta[\pi(\theta|\mathbf{x})] = \frac{n+1}{n + \sum_{i=1}^n x_i + 2}$$

1) 代入 $n=1$ 和 $x_1=3$, 得到

$$\hat{\theta}_B = \frac{1+1}{1+3+2} = \frac{1}{3}$$

2) 代入 $n=3$ 和 $x_1=2, x_2=3, x_3=5$, 得到

$$\hat{\theta}_B = \frac{3+1}{3+10+2} = \frac{4}{15}$$

□