期中考试

1. (10') 设 $X_1,\cdots,X_n(n>6)$ 是来自指数分布 $p(x,\lambda)=\lambda e^{-\lambda x}I(x\geqslant 0)$ 的样本，其中 $\lambda>0$。用 $X_{(i)}$ 表示样本的第 i 个次序统计量。求解：

   (a) λ 的充分统计量 T；

   解：

   样本联合密度为

   $$p(x_1,x_2,\cdots;\lambda)=\lambda^n e^{-n\bar{x}\lambda}I(x_{(1)}\geqslant 0)$$

   由由因子分解定理知，$T=\bar{x}$ 为 λ 的充分统计量。 □

   (b) $\mathrm{Cov}(X_{(3)},X_{(6)})$。

   解：

   设 $Y=\lambda X$，则 $Y_1,Y_2,\cdots,Y_n$ 为来自 $\mathrm{Exp}(1)$ 的样本，所以 $\mathrm{Cov}(X_{(3)},X_{(6)})=\frac{1}{\lambda^2}\mathrm{Cov}(Y_{(3)},Y_{(6)})$，下面只需求三个量：$E(Y_{(3)}Y_{(6)})$，$E(Y_{(3)})$，$E(Y_{(6)})$。

   $$\begin{aligned}\because p_3(u)&=\frac{n!}{2!(n-3)!}F^2(u)[1-F(u)]^{n-3}p(u)\\&=\frac{n!}{2!(n-3)!}(1-e^{-u})^2e^{-(n-2)u}I(u\geqslant 0)\\\therefore E(Y_{(3)})&=\int_0^{+\infty}up_3(u)\mathrm{d}u=\int_0^{+\infty}\frac{n!}{2!(n-3)!}u(1-e^{-u})^2e^{-(n-2)u}\mathrm{d}u\\&=\frac{n^4-3n^2+6n-2}{n(n-1)(n-2)}\end{aligned}$$

   同理算出 $E(Y_{(6)})$ 与 $E(Y_{(3)}Y_{(6)})$ 后可计算出

   $$\mathrm{Cov}(X_{(3)},X_{(6)})=\frac{1}{\lambda^2}\mathrm{Cov}(Y_{(3)},Y_6)=\frac{1}{\lambda^2}\left[E(Y_{(3)}Y_{(6)})-E(Y_{(3)})E(Y_{(6)})\right]$$

   □

2. (15') 设 X_i，$i=1,2,3$ 独立且都分别服从 $N(i,i^2)$，利用 X_i 构造下面分布的统计量：

   (a) 自由度为 3 的 χ^2 分布；

   $$\left(\frac{X_1-1}{1}\right)^2+\left(\frac{X_2-2}{2}\right)^2+\left(\frac{X_3-3}{3}\right)^2=\chi^2(3)$$

   (b) 自由度为 2 的 t 分布；

$$\frac{(X_1-1)\sqrt{2}}{\sqrt{\left(\frac{X_2-2}{2}\right)^2+\left(\frac{X_3-3}{3}\right)^2}} \sim t(2)$$

(c) 自由度为 1, 2 的 F 分布。

$$\frac{\left(\frac{X_1-1}{1}\right)^2}{\left(\frac{X_2-2}{2}\right)^2+\left(\frac{X_3-3}{3}\right)^2} \sim F(1,2)$$

3. (10') 设 $X_1,\cdots,X_n$ 是来自均值为 μ，方差为 σ^2 的总体的简单样本。

(a) 证明：如果 $\sum_{i=1}^n a_i=1$，则估计量 $\sum_{i=1}^n a_iX_i$ 是 μ 的一个无偏估计量；

证明：

$$E\left(\sum_{i=1}^n a_iX_i\right)=\sum_{i=1}^n a_i(EX_i)=\left(\sum_{i=1}^n a_i\right)\mu=1\cdot\mu=\mu$$

所以估计量 $\sum_{i=1}^n a_iX_i$ 是 μ 的一个无偏估计量。 □

(b) 在所有这类形式的估计量中求一个最小方差者，并计算其方差。

解：

$$\operatorname{Var}\left(\sum_{i=1}^n a_iX_i\right)=\left(\sum_{i=1}^n a_i^2\right)\sigma^2=\left(\sum_{i=1}^n a_i^2\cdot\sum_{i=1}^n 1^2\right)\sigma^2\frac{1}{n}\geqslant\left(\sum_{i=1}^n a_i\right)^2\sigma^2$$

所以当 $\sum_{i=1}^n a_i=1$ 时方差最小，为 σ^2。 □

4. (15') 设简单样本 $X_1,\cdots,X_n\sim F$。对给定常数 x_0，令 $F_n(x_0)=\frac{1}{n}\sum_{i=1}^n I(X_i\leqslant x_0)$。回答以下问题：

(a) 证明 $F_n(x_0)$ 是 $F(x_0)$ 的无偏估计；

证明：

$$\begin{aligned}&\because \sum_{i=1}^n I(x_i\leqslant x_0)\text{为}n\text{重伯努利的独立和形式}\\&\therefore E\left(\sum_{i=1}^n I(x_i\leqslant x_0)\right)=nE(I(x_1\leqslant x_0))=nF(x_0)\end{aligned}$$

□

(b) 证明 $F_n(x_0)$ 是 $F(x_0)$ 的相合估计;

证明:

由于 $E(F_n(x_0)) = F(x_0)$, $\mathrm{Var}(F_n(x_0)) = \frac{1}{n}F_n(x_0)(1 \cdot F(x_0)) \to 0$ $(n \to \infty)$, 因此 $F_n(x_0)$ 是 $F(x_0)$ 的相合估计。 □

(c) 证明 $F_n(x_0)$ 的渐近正态性。

证明:

$$\because F_n(x_0)\text{为独立和}$$
$$\therefore \text{由中心极限定理知}\sqrt{n}(F_n(x_0) - F(x_0)) \xrightarrow{L} N(0, F(x_0)(1 - F(x_0)))$$

□

5. (15') 设随机变量 $Y_1, \cdots, Y_n$ 满足

$$Y_i = x_i\beta + \varepsilon_i,\ i = 1, \cdots, n,$$

其中 $x_1, \cdots, x_n$ 是固定常数, $\varepsilon_1, \cdots, \varepsilon_n$ 独立同分布于 $N(0, \sigma^2)$, 其中 σ^2 未知。回答以下问题:

(a) 求关于 (β, σ^2) 的一个 2 维充分统计量。

解:

由于 $Y_1, Y_2, \cdots, Y_n$ 的联合密度函数为 $Y_1 \sim N(x_i\beta, \alpha^2)$,

$$\begin{aligned}
L &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(Y_i - x_i\beta)^2} \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - x_i\beta)^2\right\} \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} Y_i^2} \cdot e^{\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2x_iY_i\beta} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\beta^2}
\end{aligned}$$

设 $T_1 = \sum_{i=1}^{n} Y_i^2$, $T_2 = \sum_{i=1}^{n} x_iY_i$, 有因子分解定理知 (T_1, T_2) 为 (β, α^2) 的二维充分统计量。 □

(b) 求 β 的 MLE 并证明它是 β 的一个无偏估计;

令 $\frac{\mathrm{d}L}{\mathrm{d}\beta} = 0$, 得 $-\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - x_i\beta)x_i = 0 \implies \hat{\beta}_{ML} = \frac{\sum_{i=1}^{n} Y_ix_i}{\sum_{i=1}^{x} x_i^2} = \frac{\overline{Yx}}{\overline{x^2}}$。

$$E\left(\hat{\beta}_{ML}\right)=E\left(\frac{1}{\sum_{i=1}^{n}x_i^2}\sum_{i=1}^{n}x_iY_i\right)=\frac{1}{\sum_{i=1}^{n}x_i^2}\sum_{i=1}^{n}E(x_iY_i)=\frac{1}{\sum_{i=1}^{n}x_i^2}\sum_{i=1}^{n}x_i(EY_i)$$
$$=\frac{1}{\sum_{i=1}^{n}x_i^2}\sum_{i=1}^{n}x_i(x_i\beta)=\beta$$

(c) 求 β 的 MLE 的分布。

解：

$$\hat{\beta}_{ML}=\frac{\sum_{i=1}^{n}x_iY_i}{\sum_{i=1}^{n}x_i^2}$$

□

6. (15') 设 $X_1,X_2,\cdots,X_n$ 是来自 $N(\mu,1)$ 的简单随机样本，其中 μ 未知。用 $\bar{X}$ 表示样本均值。令 $p=P(X_1\geqslant 0)$，回答以下问题:

(a) 求 p 的极大似然估计 $\hat{p}$;

解：

因为 μ 的 MLE 为 $\bar{x}$，而 $p=P(X_1\geqslant 0)=P\left(\frac{x_1-\mu}{1}\geqslant -\mu\right)=1-\Phi(-\mu)=\Phi(\mu)$，所以由 MLE 的不变性知 $\hat{P}_{MLE}=\Phi(\bar{x})$ □

(b) 求 $\sqrt{n}(\hat{p}-p)$ 的极限分布;

解：

因为 $\hat{\mu}=\bar{x}\sim N(\mu,\frac{1}{n})$，设 $g(x)=\Phi(x)$，由 Delta 方法知

$$\sqrt{n}\,(\hat{p}-p)=\sqrt{n}\,(\Phi(\bar{x})-\Phi(\mu))\xrightarrow{L}N(0,\phi^2(\mu)\frac{1}{n})$$

其中 ϕ 为标准正态分布的概率密度函数。 □

(c) 求 p 的 UMVUE。

解：

$\Phi\left(\sqrt{\frac{n}{n-1}}\cdot\bar{x}\right)$ 为 p 的 UMVME。 □

7. (10') 已知总体密度为 $p(x;\theta)=\theta^2x^{\theta^2-1}(\theta>0,0<x<1)$。若有容量为 n 的样本，求参数 θ 无偏估计的 C-R 下界。

解：

$$
\begin{aligned}
\because \quad & \ln p = 2\ln\theta + (\theta^2 - 1)\ln x \\
& \frac{\partial \ln p}{\partial \theta} = \frac{2}{\theta} + 2\theta \ln x \\
& \frac{\partial^2 \ln p}{\partial \theta^2} = -\frac{2}{\theta^2} + 2\ln x \\
\therefore \quad & I(\theta) = -E\left(\frac{\partial^2 \ln p}{\partial \theta^2}\right) = \frac{2}{\theta^2} - 2E(\ln x) \\
& \quad = \frac{2}{\theta^2} - 2\int_0^1 \ln x \cdot \theta^2 x^{\theta^2-1} \mathrm{d}x = \frac{2}{\theta^2} + \frac{2}{\theta^2} = \frac{4}{\theta^2} \\
\therefore \quad & \quad \text{C-R 下界为} \frac{1}{nI(\theta)} = \frac{\theta^2}{4n}
\end{aligned}
$$

□

8. 设 $X_1, \cdots, X_n$ 是来自泊松分布 Poisson(λ)（其中 $\lambda > 0$）的一个样本。参数 λ 的先验分布是 $\phi(\lambda) = e^{-\lambda}$（其中 $\lambda > 0$），求 λ 的后验密度和贝叶斯估计。

解：

后验分布为

$$
\begin{aligned}
h(\theta | X_1, \cdots, X_n) &= c\prod_{i=1}^{n} f(x_i|\lambda) \cdot \pi(\lambda) \\
&= \frac{\lambda^{X_1+X_2+\cdots+X_n}}{x_1! x_2! \cdots X_n!} e^{-n\lambda} e^{-\lambda} \\
&= c\lambda^{\sum_{i=1}^n x_i} e^{-(n+1)\lambda} \sim \mathrm{Ga}\left(\sum_{i=1}^{n} x_i + 1, n+1\right)
\end{aligned}
$$

后验均值为贝叶斯估计 $\hat{\lambda} = \dfrac{\sum_{i=1}^n x_i + 1}{n+1}$。 □

附加题

1. (10') 设 $X_1, \cdots X_n, X_{n+1}$ 是来自 $N(\mu, \sigma^2)$ 的样本，记 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$，$S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$。求解:

   (a) 常数 a 和 b 使得 $T = (aX_{n+1} + b\bar{X}_n)/S_n$ 服从 t 分布，并指出 t 分布的自由度;

   解：

   若 $T = (aX_{n+1} + b\bar{X}_n)/S_n$ 服从 t 分布，则因为 $X_{n+1}, \bar{X}_n$ 都与 S_n 独立，所以

$$
\begin{aligned}
ET = 0 &\implies aE(X_{n+1}) + b(\bar{X}_n) = 0 \\
&\implies a\mu + b\mu = 0 \implies a = -b
\end{aligned}
$$

又因为 $\dfrac{(\bar{X}_n - X_{n+1})}{\sqrt{\frac{n+1}{n}}\sigma} = \dfrac{\bar{X}_n - \mu - (X_{n+1} - \mu)}{\sqrt{\frac{n+1}{n}}\sigma} \sim N(0,1)$，$\dfrac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$，所以

$$T = \frac{(\bar{X}_n - X_{n+1})}{\sqrt{\frac{n+1}{n}}\sigma} \Bigg/ \sqrt{\frac{(n-1)S_n^2}{\sigma^2(n-1)}} \sim t(n-1)$$

整理得 $\sqrt{\dfrac{n}{n+1}}(\bar{X}_n - X_{n+1})/S_n \sim t(n-1)$
即 $a = \pm\sqrt{\frac{n}{n+1}}, b = \mp\sqrt{\frac{n}{n+1}}$，自由度为 $n-1$。 □

(b) 计算 $\mathbb{E}(S_n^3)$。

解：

$$\because X = \frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$$
$$\therefore X\text{的密度函数为}p(x) = \frac{\left(\frac{1}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \ (x>0)$$

取 $y = \sqrt{x}$，则 $Y = \sqrt{X}$ 的密度函数为

$$g(y) = p(x)\frac{\mathrm{d}x}{\mathrm{d}y} = p(x)\cdot 2y = p(y^2)\cdot 2y = \frac{\left(\frac{1}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \cdot 2y^{n-2}\cdot y\cdot e^{-\frac{y^2}{2}}$$

$$\begin{aligned} E(Y^3) &= \int_0^{+\infty} y^3 g(y)\mathrm{d}y = \int_0^{+\infty} \frac{2\left(\frac{1}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} y^{n+1}\cdot e^{-\frac{y^2}{2}}\mathrm{d}y \\ &= \frac{2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^{+\infty} y^{n+1} e^{-\frac{y^2}{2}}\mathrm{d}y = \begin{cases} \frac{2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}(2k-1)!!, & n = 2k-1 \\ \frac{2^{\frac{n+1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}(2k)!!, & n = 2k \end{cases} \end{aligned}$$

由 $\sqrt{\dfrac{(n-1)S_n^2}{\sigma^2}} = Y \implies ES_n^3 = \left(\dfrac{\sigma}{\sqrt{n-1}}\right)^3 \cdot E(Y^3)$ 代入即得。 □

注： 正常题总分为 100 分，附加题总分为 10 分。总评分为两者之和，但总评分不超过 100 分。