北京大学信息科学技术学院试题

考试科目: ______________________ 姓名: ________________ 学号: ________________
考试时间: _________年______月______日 任课教师: __________________

注意事项:

1. 考试时间为 120 分钟（上午 8:30 至 10:30）。
2. 共 4 个大题，总分为 100 分，试卷共**两页**。请合理分配时间，并将所有答案**写在答题纸**上，试题纸或草稿纸上作答无效。
3. 考试结束后将回收答题纸，试卷和草稿纸。

题目一（共 35 分）

总体 X 服从概率密度函数满足

$$f(x) = \begin{cases} \frac{2\theta^2}{x^3} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

的连续分布，其中 $\theta > 0$ 为未知参数。令 $X_1, X_2, \ldots, X_n$ 为简单随机样本。

1. (5 分）给出 X 的分布函数 $F(x) = P(X \leq x)$。
2. (5 分）计算 X 的数学期望 $E(X)$。将未知参数 θ 写为 $E(X)$ 的函数，并将 $E(X)$ 替换为样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$，给出 θ 的矩法估计量 $\hat{\theta}_1$。判断 $\hat{\theta}_1$ 是否为无偏估计量和渐进无偏估计量。
3. (5 分）给出未知参数 θ 的最大似然估计量 $\hat{\theta}_2$。
4. (10 分）判断 $\hat{\theta}_2$ 是否为无偏估计量和渐进无偏估计量。当 $n \geq 2$ 时，给出 $\hat{\theta}_2$ 的均方误差。判断 $\hat{\theta}_2$ 是否为一致估计量。
5. (5 分）利用 $\hat{\theta}_2$，给出 $\hat{\theta}_L$ 使得 $P(\hat{\theta}_L \leq \theta) = 1 - \alpha$，也即 $\hat{\theta}_L$ 为 θ 的置信水平为 $1-\alpha$ 的单侧置信下限。
6. (5 分）考虑假设检验问题：原假设 $H_0 : \theta = \theta_0$，备择假设 $H_1 : \theta > \theta_0$。这里 $\theta_0 > 0$ 为给定常数。若拒绝域为 $W = ((x_1, x_2, \ldots, x_n) \mid \hat{\theta}_2 \geq c)$，给出显著性水平为 α 时待定值 c 的取值。

题目二（共 15 分）

1. (5 分) 给定 $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$，其中 x_i, y_i 均为实数。令 $Q_\gamma(\alpha, \beta) = \sum_{i=1}^n (y_i - \beta x_i - \alpha)^2 + \gamma\beta^2$，这里 $\gamma > 0$ 为参数。令 $\hat{\alpha}_\gamma, \hat{\beta}_\gamma = \text{argmin}_{\alpha,\beta} Q_\gamma(\alpha, \beta)$，也即 $\hat{\alpha}_\gamma$ 和 $\hat{\beta}_\gamma$ 为使 $Q_\gamma(\alpha, \beta)$ 最小化的 α 和 β。仿照最小二乘估计，写出正规方程，并给出 $\hat{\alpha}_\gamma$ 和 $\hat{\beta}_\gamma$ 的表达式。
2. (10 分) 假设 $y_i = \beta x_i + \alpha + \epsilon_i$。测量误差（或噪声）$\epsilon_i$ 为随机变量，满足 $E(\epsilon_i) = 0, \text{Var}(\epsilon_i) = \sigma^2$，且 ϵ_i 相互独立。这里 α, β 和 $\sigma^2 > 0$ 为未知参数。判断 $\hat{\alpha}_\gamma$ 和 $\hat{\beta}_\gamma$ 是否为 α 和 β 的无偏估计量，并给出 $\hat{\alpha}_\gamma$ 和 $\hat{\beta}_\gamma$ 的均方误差。

题目三（共 30 分）

二维连续随机变量 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，其联合密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\cdot\left(\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)\right]。$$

这里 $\sigma_1,\sigma_2>0$，$|\rho|<1$。令 $U=e^X,V=e^Y$。

1. （10 分）计算 $E(U),E(V),\mathrm{Var}(U),\mathrm{Var}(V)$ 和 $\mathrm{Cov}(U,V)$，并给出 U 和 V 的相关系数。
2. （10 分）给出 (U,V) 的联合密度函数，以及 U 和 V 各自的边际密度函数。
3. （10 分）对于 $v>0$，给出给定 $V=v$ 条件下 U 的条件密度函数，并计算 U 的条件数学期望。

题目四（共 20 分）

$X_1,X_2,\ldots,X_n$ 为独立同分布的随机变量，X_i 服从柯西分布，其概率密度函数满足对于任意实数 x，$f(x)=\frac{1}{\pi(x^2+1)}$。令 $Y_1=|X_1|,Y_2=|X_2|,\ldots,Y_n=|X_n|$。

1. （5 分）给出 Y_i 的边际密度函数。
2. （5 分）证明存在常数 $c_1>0$，
$$P\left(\sum_{i=1}^{n}Y_n\le c_1n^2\right)\ge 2/3。$$
3. （5 分）证明存在常数 $c_2>0$，
$$P\left(\sum_{i=1}^{n}Y_n\ge c_2n\right)\ge 2/3。$$
4. （5 分）给出函数 $f(n)$，使得存在常数 $c_3,c_4>0$，对于充分大的 n，
$$P\left(\sum_{i=1}^{n}Y_n\le c_3f(n)\right)\ge 2/3$$
且
$$P\left(\sum_{i=1}^{n}Y_n\ge c_4f(n)\right)\ge 2/3。$$
给出比第二问和第三问中更好的上界或下界均可得到部分分。