信息学中的概率统计：作业五

截止日期：2025 年 11 月 26 日（周三）下午 3 点前。请务必通过教学网提交电子版。可下课前同时提交纸质版。

与本次作业相关的事宜,请发邮件给我(ruosongwang@pku.edu.cn),抄送研究生助教叶昊洋(yhyfhgs@gmail.com),以及负责本次作业的本科生助教疏彦凯（syksykccc@stu.pku.edu.cn）。

第一题

随机变量 $X_1, X_2, \ldots, X_n$ 相互独立，且 $X_i \sim \mathrm{Exp}(\lambda_i)$，这里对于任意 $1 \le i \le n$，$\lambda_i > 0$。给出 $Y = \min\{X_1, X_2, \ldots, X_n\}$ 服从的分布。

第二题

随机变量 $X \sim U(0,1)$，$Y \sim U(0,1)$。且 X 与 Y 相互独立。令 $U = X + Y$，$V = X - Y$。

(1) 使用卷积公式，计算 U 和 V 的边际密度函数。

(2) 使用密度变换公式，计算 (U, V) 的联合密度函数。

(3) 利用上一问中的联合密度函数，重新计算 U 和 V 的边际密度函数。判断 U 与 V 是否相互独立。

(4) 对于任意 $0 < u < 2$，计算给定 $U = u$ 条件下 V 的条件密度函数。

(5) 计算 $\mathrm{Corr}(U, V)$，并判断 U 与 V 的相关性。

第三题（选做）

在课上，我们证明了如下结论：X 服从 n 维高斯分布 $N(\mu, B)$，令 $Y = AX + b$，若 $A \in \mathbb{R}^{n\times n}$ 可逆，则 $AX + b \sim N(A\mu + b, ABA^\top)$。本题中，我们将证明当 $A \in \mathbb{R}^{m\times n}$ 为行满秩矩阵（$\mathrm{rank}(A) = m < n$）时，结论同样成立。

注意：在本题中请不要使用服从高维高斯分布的随机向量的子向量仍服从高斯分布的结论，因为该结论的证明需要用到本题的结论。

(1) 证明存在矩阵 $Q \in \mathbb{R}^{m\times n}$ 和 $R \in \mathbb{R}^{m\times m}$，使得

  (a) $R \in \mathbb{R}^{m\times m}$ 为可逆矩阵；

  (b) Q 的每一行的均为单位向量（模长均为 1），且不同行正交；

  (c) $A = RQ$。

(2) 证明存在正交矩阵 $\tilde{Q}$，其前 m 行与 Q 相同。

(3) 若 $Z \sim N(0, I_n)$，也即独立同分布的标准高斯分布，给出 $\tilde{Q}Z$ 服从的分布，并基于边际密度函数的公式，给出 QZ 服从的分布。

(4) 使用上一问中的结论，若 $Z \sim N(0, I_n)$，给出 $AZ + b$ 服从的分布。这里 $A \in \mathbb{R}^{m\times n}$ 为行满秩矩阵。

(5) 使用上一问中的结论，若 $X \sim N(\mu, B)$，给出 $AX + b$ 服从的分布。这里同样考虑 $A \in \mathbb{R}^{m\times n}$ 为行满秩矩阵的情况。提示：证明 $AB^{1/2}$ 行满秩。

注意：本题为选做题，是否提交不会影响本次作业成绩。

第四题

令 $G = (G_1, G_2, \ldots, G_n)$ 为独立同分布的随机变量，且 $G_i \sim N(0, 1)$。对于固定的向量 $a, b \in \mathbb{R}^n$，满足 $\|a\| = \|b\| = 1$，$a \neq b$ 且 $a \neq -b$，令随机变量 $X = \sum_{i=1}^n G_i a_i$，$Y = \sum_{i=1}^n G_i b_i$。

(1) 令 $\rho = a^\top b = \sum_{i=1}^n a_i b_i$，给出 (X, Y) 服从的分布。

(2) 令 $x = r \cdot (\sqrt{1-\rho^2} \cdot \cos\theta + \rho \cdot \sin\theta), y = r\sin\theta$。验证 $x^2 + y^2 - 2\rho xy = r^2(1-\rho^2)$。

(3) 令二维连续随机变量 (R, Θ) 满足 $R \geq 0, \Theta \in [0, 2\pi]$，且

$$\begin{cases} X = R \cdot (\sqrt{1-\rho^2} \cdot \cos\Theta + \rho \cdot \sin\Theta) = R \cdot \sin(\arccos\rho + \Theta) \\ Y = R\sin\Theta \end{cases}。$$

计算 R, Θ 的联合密度函数，R 和 Θ 的边际密度函数，并判断 R 和 Θ 的独立性。

(4) 计算 $P(X \geq 0, Y \geq 0)$。

(5) 若对于一般的向量 $a, b \in \mathbb{R}^n$，（不要求单位向量，也不排除 $a = b$ 或 $a = -b$ 的情况），重新计算 $P(X \geq 0, Y \geq 0)$。

第五题

令 $X_1, X_2, \ldots, X_n$ 为独立同分布随机变量，且 $X_i \sim N(0, 1)$。令 $Y = \sum_{i=1}^n X_i^2$。

(1) 给出 Y 服从的分布。提示：参考作业三第六题。

(2) 对于任意实数 $t < 1/4$，证明

$$E(e^{t(Y-n)}) \leq e^{2t^2 n}。$$

提示：参考作业零中的不等式。

(3) 对于任意 $0 \leq \Delta < 1$，证明

$$P(Y \geq (1+\Delta)n) \leq e^{-n\Delta^2/8}。$$

提示：根据 $0 \leq \Delta < 1$，选择合适的 t，并使用马尔可夫不等式。

(4) 对于任意 $0 \leq \Delta < 1$，证明

$$P(Y \leq (1-\Delta)n) \leq e^{-n\Delta^2/8}。$$

第六题

(1) 在课上我们考虑了如下矩阵 $A \in \mathbb{R}^{n\times n}$：对于任意 $1 \leq i, j \leq n$，$A_{i,j} \sim N(0, 1)$，且不同元素相互独立。计算 $E(\mathrm{trace}(A^3))$ 和 $E(\mathrm{trace}(A^4))$。提示：首先考虑 $n = 1$ 的情况，并参考第五题。

(2) 考虑如下对称矩阵 $B \in \mathbb{R}^{n\times n}$：对于任意 $1 \leq i \leq j \leq n$，$B_{i,j} \sim N(0, 1)$，且不同元素相互独立；当 $i > j$，有 $B_{i,j} = B_{j,i}$。计算 $E(\mathrm{trace}(B^2))$ 和 $E(\mathrm{trace}(B^4))$。

(3) 令 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为上一问中对称矩阵 B 的特征值。证明

$$P\left(\max_{1\leq i\leq n} |\lambda_i| \leq 12306 n^{3/4}\right) \geq 0.9。$$