SI 140A-02 Probability & Statistics for EECS, Fall 2024
Homework 10

Name:
Student ID:

Due on Dec. 17, 2024, 11:59 UTC+8

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points of the problem.

- Write your homework in English; otherwise you will get no points of this homework.

- Any form of plagiarism will lead to 0 point of this homework.

Problem 1

Let $U_i \sim \text{Unif}(0,1), i \geq 1$ be i.i.d. random variables. Define N as follows:

$$N = \max\left\{n : \prod_{i=1}^{n} U_i \geq e^{-1}\right\}$$

(a) Estimate $\text{E}(N)$ by generating 5000 samples of N and then use the sample mean.

(b) Estimate $\text{Var}(N)$.

(c) Estimate $P(N = i)$, for $i = 0, 1, 2, 3$.

(d) Can you find the exact distribution of N?

Problem 2

Visualization of Bivariate Normal Distribution with correlation coefficient ρ

(a) Use the following transformation to generate samples from bivariate Normal distribution with correlation coefficient ρ :

$$X = Z$$
$$Y = \rho Z + \sqrt{1 - \rho^2} W$$

where $-1 < \rho < 1$, Z and W are i.i.d. random variables satisfying $\mathcal{N}(0,1)$.

(b) Plot the joint pdf function and the corresponding contour(or isocontour) as the following figure:

(a) $\rho = -0.8$ (b) $\rho = -0.4$ (c) $\rho = 0$ (d) $\rho = 0.4$ (e) $\rho = 0.8$
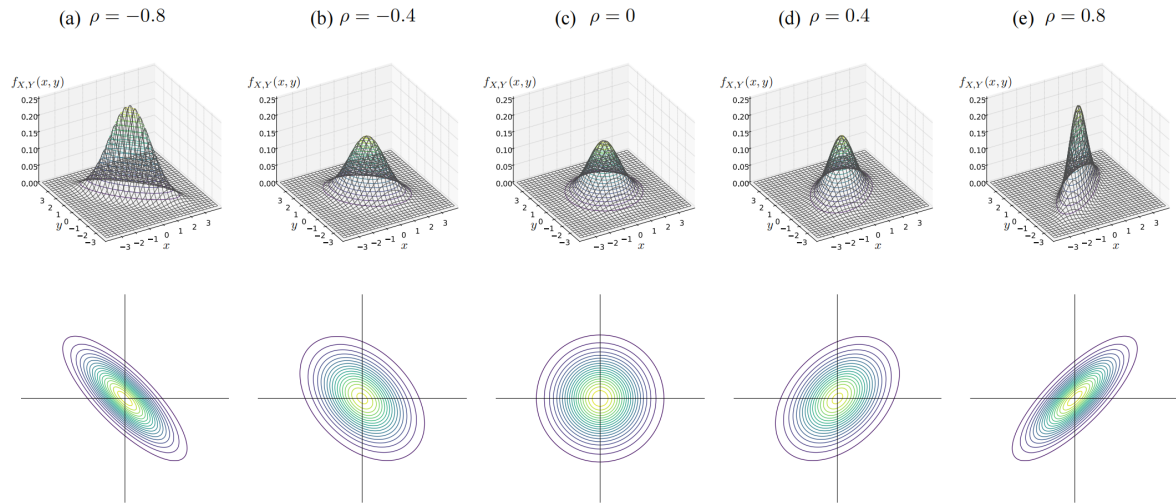
Problem 3

Let $X_1 \sim \text{Expo}\left(\lambda_1\right)$, $X_2 \sim \text{Expo}\left(\lambda_2\right)$ and $X_3 \sim \text{Expo}\left(\lambda_3\right)$ be independent.

(a) Find $E\left(X_1 \mid X_1 > 2024\right)$

(b) Find $E\left(X_1 \mid X_1 < 1997\right)$

(c) Find $E\left(X_1 + X_2 + X_3 \mid X_1 > 1997, X_2 > 2014, X_3 > 2025\right)$ in terms of $\lambda_1, \lambda_2, \lambda_3$.

Problem 4

Let X and Y be two continuous random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6xy & \text{if } 0 \leq x \leq 1, 0 \leq y \leq \sqrt{x} \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the marginal distributions of X and Y. Are X and Y independent?

(b) Find $E[X \mid Y = y]$ and $\text{Var}[X \mid Y = y]$ for $0 \leq y \leq 1$.

(c) Find $E[X \mid Y]$ and $\text{Var}[X \mid Y]$.

Problem 5

Let X be a discrete r.v. whose distinct possible values are $x_0, x_1, \ldots$, and let $p_k = P(X = x_k)$. The entropy of X is $H(X) = \sum_{k=0}^{\infty} p_k \log_2 (1/p_k)$.

(a) Find $H(X)$ for $X \sim \text{Geom}(p)$.

(b) Let X and Y be i.i.d. discrete r.v.s. Show that $P(X = Y) \geq 2^{-H(X)}$.

Problem 6

Instead of predicting a single value for the parameter, we give an interval that is likely to contain the parameter: A $1 - \delta$ confidence interval for a parameter p is an interval $[\hat{p} - \epsilon, \hat{p} + \epsilon]$ such that $Pr(p \in [\hat{p} - \epsilon, \hat{p} + \epsilon]) \geq 1 - \delta$. Now we toss a coin with probability p landing heads and probability $1 - p$ landing tails. The parameter p is unknown and we need to estimate its value from experiment results. We toss such coin N times. Let $X_i = 1$ if the i th result is head, otherwise 0 . We estimate p by using

$$\hat{p} = \frac{X_1 + \ldots + X_N}{N}$$

Find the $1 - \delta$ confidence interval for p, then discuss the impacts of δ and N.

(a) Method 1: Adopt Chebyshev inequality to find the $1 - \delta$ confidence interval for p, then discuss the impacts of δ and N.

(b) Method 2: Adopt Hoeffding bound to find the $1 - \delta$ confidence interval for p, then discuss the impacts of δ and N.

(c) Discuss the pros and cons of the above two methods.

Problem 7

(**Optional Challenging Problem**) We consider the progressive Monty Hall problem. This time we assume there are n identical doors, where n is an integer satisfying $n \geq 3$. One door conceals a car, the other $n - 1$ doors conceal goats. You choose one of the doors at random but do not open it. Monty then opens a door he knows to conceal a goat, always choosing randomly among the available doors. At this point he gives you the option either of sticking with your original door or switching to one of the remaining doors. You make your decision. Monty now eliminates another goat-concealing door (at random) and once more gives you the choice either of sticking or switching. This process continues until only two doors remain in play. What strategy should you follow to maximize your chances of winning? We consider three strategies: (1) Select a door at random and stick with it throughout. (2) Select a door at random, then switch doors at every opportunity, choosing your door randomly at each step. (3) Select a door at random, stick with your first choice until only two doors remain, and then switch. When $n = 10$ and $n = 1000$, **please run simulations to estimate the winning probability of each strategy. Check which strategy is best and provide the corresponding intuition.**