Name:
Roll Number:
Department:
Program: BTech / MTech TA / MTech RA / PhD (Tick one)

CS6660: Mathematical Foundations of Data Science

Mid Term Exam 1

Date: 06 October 2024

Instructions:

- **This exam is for a total of 30 MARKS.**
- **You are allowed to keep ONE A4 sheet of written material containing formulae.**
- **Hints are provided for some questions.**
  However, it is NOT mandatory to solve the question using the approach in the hints.
  If you think you have a better approach in mind than the one given in the hint, feel free to present your approach.
- **Show all your working clearly.**
  We want to see your thought process, and possibly provide partial credit for the intermediate logical steps.
- **Plagiarism will NOT be entertained at any length.**
  If you are caught cheating during the exam, your answer script will NOT be evaluated.

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Assume that all random variables appearing in the questions below are defined with respect to $\mathscr{F}$.
Assume that all logarithms appearing in the questions below are natural logarithms, unless explicitly stated otherwise.

Name:
Roll Number:
Department:
Program: BTech / MTech TA / MTech RA / PhD (Tick one)

1. (a) **(4 Marks)**
Fix $M \in \mathbb{N}$.
Let $\Omega_1, \ldots, \Omega_M$ be a *partition* of Ω, i.e., $\Omega_i \cap \Omega_j = \emptyset$ for all $i \neq j$, and $\bigcup_{i=1}^{M} \Omega_i = \Omega$.
Let A and B be two events such that
   - A and B are conditionally independent conditioned on Ω_i, for all $i \in \{1, \ldots, M\}$, and
   - B is independent of Ω_i for all $i \in \{1, \ldots, M\}$.

   Prove that A and B are independent.

   (b) **(1 Mark)**
A box contains three coins: two regular coins and one fake 2-headed coin.
A friend of yours picks a coin uniformly at random, tosses it, and tells you that it landed heads up.
What is the probability that your friend chose the 2-headed coin?

Solution: We present the solution to each of the parts below.

(a) We are given that

$$\mathbb{P}(A \cap B|\Omega_i) = \mathbb{P}(A|\Omega_i) \cdot \mathbb{P}(B|\Omega_i), \qquad \mathbb{P}(B|\Omega_i) = \mathbb{P}(B), \qquad \forall i \in \{1, \ldots, n\}.$$

Using the above fact along with the law of total probability, we have

$$\mathbb{P}(A \cap B) = \sum_{i=1}^{n} \mathbb{P}(A \cap B|\Omega_i) \cdot \mathbb{P}(\Omega_i) = \sum_{i=1}^{n} \mathbb{P}(A|\Omega_i) \cdot \mathbb{P}(B) \cdot \mathbb{P}(\Omega_i) = \mathbb{P}(B) \cdot \sum_{i=1}^{n} \mathbb{P}(A|\Omega_i) \cdot \mathbb{P}(\Omega_i) = \mathbb{P}(B) \cdot \mathbb{P}(A),$$

thus proving that $A \perp\!\!\!\perp B$.

(b) We present two solutions, assuming that the two regular coins are identical vs distinct.

- <u>Assuming that the two regular coins are identical:</u>
  Let A, B denote respectively the event of choosing regular coin and choosing fake coin. Let E denote the event of occurrence of head. Then, we have
  $$\mathbb{P}(E|A) = \frac{1}{2}, \qquad \mathbb{P}(E|B) = 1.$$
  Furthermore, we have $\mathbb{P}(A) = 2/3 = 1 - \mathbb{P}(B)$. Our interest is to compute the value of $\mathbb{P}(B|E)$. Using Bayes' Theorem, we have
  $$\mathbb{P}(B|E) = \frac{\mathbb{P}(E|B) \cdot \mathbb{P}(B)}{\mathbb{P}(E|A) \cdot \mathbb{P}(A) + \mathbb{P}(E|B) \cdot \mathbb{P}(B)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{3} \cdot 1} = \frac{1}{2}.$$
- <u>Assuming that the two regular coins are distinct:</u>
  Let A, B, C denote respectively the event of choosing regular 1, choosing regular coin 2, and choosing choosing fake coin. Let E denote the event of occurrence of head. Then, we have
  $$\mathbb{P}(E|A) = \frac{1}{2}, \qquad \mathbb{P}(E|B) = \frac{1}{2}, \qquad \mathbb{P}(E|C) = 1.$$
  Furthermore, we have $\mathbb{P}(A) = 1/3 = \mathbb{P}(B) = \mathbb{P}(C)$. Our interest is to compute the value of $\mathbb{P}(C|E)$. Using Bayes' Theorem, we have
  $$\mathbb{P}(C|E) = \frac{\mathbb{P}(E|C) \cdot \mathbb{P}(C)}{\mathbb{P}(E|A) \cdot \mathbb{P}(A) + \mathbb{P}(E|B) \cdot \mathbb{P}(B) + \mathbb{P}(E|C) \cdot \mathbb{P}(C)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} \cdot 1} = \frac{1}{2}.$$

2. (a) **(2 Marks)**
Two numbers are drawn independently and uniformly from the unit interval $[0, 1]$.
The smaller of the two the numbers is known to be less than $1/3$.
What is the probability that the larger one is greater than $3/4$?

(b) **(3 Marks)**
A miner is trapped in a mine containing three doors. The first door leads to a tunnel that takes him to safety after two hours of travel. The second door leads to a tunnel that returns him to the mine after three hours of travel. The third door leads to a tunnel that returns him to the mine after five hours. Assuming that the miner is at all times equally likely to choose any one of the three doors, what is the expected length of time until the miner reaches safety?
Hint: Let D be a random variable that denotes the door chosen by the miner, and let T denote the time until safety. Compute $\mathbb{E}[T|\{D = 1\}]$, $\mathbb{E}[T|\{D = 2\}]$, and $\mathbb{E}[T|\{D = 3\}]$ separately, and use the law of iterated expectations to compute $\mathbb{E}[T]$.

Solution: We present the solution to each of the parts below.

(a) Let X_1 and X_2 denote the two numbers. The desired probability is

$$\mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\}\middle|\left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\}\right) = \frac{\mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\} \cap \left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\}\right)}{\mathbb{P}\left(\left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\}\right)}.$$

The numerator of the above expression may be computed, via the law of total probability, as follows:

$$\begin{aligned}
&\mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\} \cap \left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\}\right) \\
&= \mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\} \cap \left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\} \cap \{X_1 \geq X_2\}\right) \\
&\qquad\qquad + \mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\} \cap \left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\} \cap \{X_1 < X_2\}\right) \\
&= \mathbb{P}\left(\left\{X_1 > \frac{3}{4}\right\} \cap \left\{X_2 < \frac{1}{3}\right\} \cap \{X_1 \geq X_2\}\right) \\
&\qquad\qquad + \mathbb{P}\left(\left\{X_2 > \frac{3}{4}\right\} \cap \left\{X_1 < \frac{1}{3}\right\} \cap \{X_1 < X_2\}\right) \\
&\overset{(a)}{=} \mathbb{P}\left(\left\{X_1 > \frac{3}{4}\right\} \cap \left\{X_2 < \frac{1}{3}\right\}\right) + \mathbb{P}\left(\left\{X_2 > \frac{3}{4}\right\} \cap \left\{X_1 < \frac{1}{3}\right\}\right) \\
&\overset{(b)}{=} 2 \cdot \mathbb{P}\left(\left\{X_1 > \frac{3}{4}\right\} \cap \left\{X_2 < \frac{1}{3}\right\}\right) \\
&\overset{(c)}{=} 2 \cdot \mathbb{P}\left(\left\{X_1 > \frac{3}{4}\right\}\right) \cdot \mathbb{P}\left(\left\{X_2 < \frac{1}{3}\right\}\right) \\
&= 2 \cdot \frac{1}{4} \cdot \frac{1}{3} \\
&= \frac{1}{6},
\end{aligned}$$

where (a) follows from the fact that $X_1 > \frac{3}{4}$ and $X_2 < \frac{1}{3}$ implies $X_1 \geq X_2$ (similarly for the other term in (a)), (b) above follows from symmetry, and (c) above follows because $X_1 \perp\!\!\!\perp X_2$.
The denominator probability term is given by

$$\mathbb{P}\left(\text{min}\{X_1, X_2\} < \frac{1}{3}\right) = 1 - \mathbb{P}\left(\left\{\text{min}\{X_1, X_2\} \geq \frac{1}{3}\right\}\right)$$

$$\begin{aligned}
&= 1 - \mathbb{P}\left(\left\{X_1 \geq \frac{1}{3}\right\} \cap \left\{X_2 \geq \frac{1}{3}\right\}\right) \\
&= 1 - \mathbb{P}(\{X_1 \geq 1/3\}) \cdot \mathbb{P}(\{X_2 \geq 1/3\}) \\
&= 1 - \frac{2}{3} \cdot \frac{2}{3} \\
&= \frac{5}{9}.
\end{aligned}$$

The desired probability is then given by

$$\mathbb{P}\left(\left\{\text{max}\{X_1, X_2\} > \frac{3}{4}\right\} \middle| \left\{\text{min}\{X_1, X_2\} < \frac{1}{3}\right\}\right) = \frac{1/6}{5/9} = \frac{3}{10}.$$

(b) The question specifies that

$$\mathbb{P}(\{D = 1\}) = \frac{1}{3} = \mathbb{P}(\{D = 2\}) = \mathbb{P}(\{D = 3\}).$$

Also, we note that

$$\mathbb{E}[T|\{D = 1\}] = 2, \qquad \mathbb{E}[T|\{D = 2\}] = 3 + \mathbb{E}[T], \qquad \mathbb{E}[T|\{D = 3\}] = 5 + \mathbb{E}[T].$$

Using the law of iterated expectations, we have

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[\mathbb{E}[T|D]] \\
&= \mathbb{E}[T|\{D = 1\}] \cdot \mathbb{P}(\{D = 1\}) + \mathbb{E}[T|\{D = 2\}] \cdot \mathbb{P}(\{D = 2\}) + \mathbb{E}[T|\{D = 3\}] \cdot \mathbb{P}(\{D = 3\}) \\
&= \frac{10}{3} + \frac{2\,\mathbb{E}[T]}{3},
\end{aligned}$$

from which it follows that $\mathbb{E}[T] = 10$.

Name:
Roll Number:
Department:
Program: BTech / MTech TA / MTech RA / PhD (Tick one)

3. Let X and Y be jointly continuous random variables with the joint PDF

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

(a) **(1 Mark)**
Plot and shade the region of integration in 2-dimensions.

(b) **(2 Marks)**
Determine the marginal PDFs of X and Y.

(c) **(1 Mark)**
Are X and Y independent? Justify your answer.

(d) **(3 Marks)**
Evaluate $\mathbb{E}[X|Y]$.

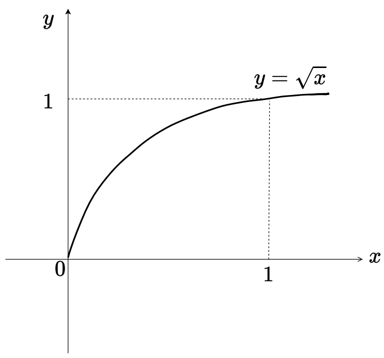
(e) **(3 Marks)**
Evaluate Var$(X|Y = 1/2)$, defined as

$$\text{Var}(X|Y = 1/2) := \mathbb{E}[X^2|\{Y = 1/2\}] - \big(\mathbb{E}[X|\{Y = 1/2\}]\big)^2.$$

You may leave the final answer in the form α/β, where α and β are co-prime (i.e., do not have any common factor).

Solution: We present the solution to each of the parts below.

(a) The region of interest is that under the curve $y = \sqrt{x}$ within the unit square (see figure below).



(b) For any $x \in [0,1]$, we have

$$f_X(x) = \int_0^{\sqrt{x}} 6xy \, \mathrm{d}y = 3x^2.$$

Also, for any $y \in [0,1]$, we have

$$f_Y(y) = \int_{y^2}^{1} 6xy \, \mathrm{d}x = 3y(1 - y^4).$$

(c) From the marginal PDF expressions, it is clear that $f_{X,Y}(1,1) = 6 \neq 0 = f_X(1)\, f_Y(1)$, thus proving that $X \not\perp Y$.

(d) To evaluate $\mathbb{E}[X|Y]$, we will first need to evaluate $f_{X|Y=y}$, the conditional PDF of X conditioned on the event $\{Y = y\}$. Noting that $f_Y(y) = 0$ for $y \in \{0, 1\}$, it follows that for any $y \in (0, 1)$,

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{6xy}{3y(1-y^4)} = \frac{2x}{1-y^4}, \qquad x \in [y^2, 1].$$

It then follows that for any $y \in (0, 1)$,

$$\mathbb{E}[X|\{Y = y\}] = \int_{y^2}^{1} x\, f_{X|Y=y}(x)\, \mathsf{d}x = \int_{y^2}^{1} \frac{2x^2}{1-y^4}\, \mathsf{d}x = \frac{2(1-y^6)}{3(1-y^4)},$$

from which it follows that

$$\mathbb{E}[X|Y] = \frac{2(1-Y^6)}{3(1-Y^4)}.$$

(e) We note that for any $y \in (0, 1)$,

$$\mathbb{E}[X^2|\{Y = y\}] = \int_{y^2}^{1} x^2\, f_{X|Y=y}(x)\, \mathsf{d}x = \int_{y^2}^{1} \frac{2x^3}{1-y^4}\, \mathsf{d}x = \frac{1-y^8}{2(1-y^4)}.$$

We then have

$$\mathsf{Var}(X|Y = 1/2) = \mathbb{E}[X^2|\{Y = 1/2\}] - \left(\mathbb{E}[X|\{Y = 1/2\}]\right)^2 = \frac{17}{32} - \left(\frac{7}{10}\right)^2 = \frac{33}{800}$$

4. On generating a random sample from the standard normal distribution on a computer.

Those of you who are familiar with Python programming language may be aware of the `numpy.random.normal()` module in Python for generating a random sample from the standard normal distribution.
In this question, we shall understand how this module works behind the scenes.
The basic premise here is that a computer first generates a random sample from Unif$(0, 1)$, and uses the uniform sample to generate the desired sample from standard normal distribution.

Let $U_1, U_2 \overset{\text{i.i.d.}}{\sim}$ Unif$(0, 1)$.

(a) **(2 Marks)**
Determine the PDF of $R = \sqrt{-2 \, \log(1 - U_1)}$.

(b) **(1 Mark)**
Let $\Theta = 2\pi U_2$. What is the distribution of Θ?

(c) **(1 Mark)**
Argue that R and Θ are independent.

(d) **(3 Marks)**
Let $X = R \, \cos(\Theta)$ and $Y = R \, \sin(\Theta)$. Show that $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$.
The module `np.random.normal()` returns only one of the two samples X and Y, and masks the other.

(e) **(1 Mark)**
Determine the variance of $Z = 3X + 4Y$.

(f) **(2 Marks)**
Let $S = X + Y$ and $D = X - Y$.
Further, let W be independent of both X and Y, with

$$\mathbb{P}(\{W = w\}) = \begin{cases} \frac{1}{2}, & w = \pm 1, \\ 0, & \text{otherwise.} \end{cases}$$

Are S and WD jointly Gaussian? Justify.

Solution: We present the solution to each of the parts below.

(a) Clearly, R is non-negative. For any $r \geq 0$,

$$\begin{aligned} \mathbb{P}(\{R \leq r\}) &= \mathbb{P}(\{\sqrt{-2 \log(1 - U_1)} \leq r\}) \\ &= \mathbb{P}(\{-2 \log(1 - U_1) \leq r^2\}) \\ &= \mathbb{P}\left(\left\{ \log(1 - U_1) \geq -\frac{r^2}{2} \right\}\right) \\ &= \mathbb{P}\left(\left\{ U_1 \leq 1 - e^{-r^2/2} \right\}\right) \\ &= 1 - e^{-r^2/2}, \end{aligned}$$

from which it follows that

$$f_R(r) = \begin{cases} re^{-r^2/2}, & r \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

The above PDF is called the *Rayleigh* PDF.

(b) We have $\Theta \sim$ Unif$(0, 2\pi)$.

(c) Because R is a function of U_1, Θ is a function of U_2, and $U_1 \perp\!\!\!\perp U_2$, it follows that $R \perp\!\!\!\perp \Theta$.

(d) Let $g_1(r,\theta) = r\cos\theta$ and $g_2(r,\theta) = r\sin\theta$. Then, we have $X = g_1(R,\Theta)$ and $Y = g_2(R,\Theta)$. Further, the the inverse of the mapping $g(r,\theta) = (g_1(r,\theta), g_2(r,\theta))$ is given by

$$g^{-1}(x,y) = (\sqrt{x^2+y^2}, \tan^{-1}(y/x)), \qquad (x,y) \in \mathbb{R}^2.$$

The Jacobian of the mapping g is given by

$$\det(J_g(r,\theta)) = \det\begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \theta} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \theta} \end{pmatrix} = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r(\cos^2\theta + \sin^2\theta) = r.$$

Setting $r = \sqrt{x^2+y^2}$ and $\theta = \tan^{-1}(y/x)$, and using the Jacobian transformation formula, we get

$$f_{X,Y}(x,y) = \frac{f_{R,\Theta}(g^{-1}(x,y))}{\left|\det(J_g(g^{-1}(x,y)))\right|} = \frac{1}{2\pi} e^{-(x^2+y^2)/2},$$

from which it follows that $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.

(e) Because X and Y are i.i.d. standard normal distributed, it follows that

$$\mathrm{Var}(3X+4Y) = \mathrm{Var}(3X) + \mathrm{Var}(4Y) = 9\,\mathrm{Var}(X) + 16\,\mathrm{Var}(Y) = 25.$$

(f) Clearly, S and D are jointly Gaussian, as X and Y are jointly Gaussian. That is, the random variable

$$c_1\,S + c_2\,D = (c_1+c_2)X + (c_1-c_2)Y, \qquad (c_1,c_2) \neq (0,0),$$

is Gaussian. We then note that

$$c_1\,S + c_2\,WD = c_1(X+Y) + c_2\,W(X-Y) = \begin{cases} (c_1+c_2)X + (c_1-c_2)Y, & \text{w.p. } \frac{1}{2}, \\ (c_1-c_2)X + (c_1+c_2)Y, & \text{w.p. } \frac{1}{2}. \end{cases}$$

Because X and Y are i.i.d. and standard normal distributed, it follows by symmetry that

$$c_1\,S + c_2\,WD \quad \overset{d}{=} \quad (c_1+c_2)X + (c_1-c_2)Y,$$

where the notation $\overset{d}{=}$ stands for "has same distribution as". In particular, we note that

$$\mathbb{E}[c_1\,S + c_2\,WD] = 0, \qquad \mathrm{Var}(c_1\,S + c_2\,WD) = (c_1+c_2)^2\,\mathrm{Var}(X) + (c_1-c_2)^2\,\mathrm{Var}(Y) = 2(c_1^2+c_2^2),$$

and therefore $c_1\,S + c_2\,WD \sim \mathcal{N}(0, 2(c_1^2+c_2^2))$. Hence, S and WD are jointly Gaussian.