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

AI5030: Probability and Stochastic Processes

Mid Term Exam 1

Date: 25 September 2024

Question	Marks Scored
1(a)	
1(b)	
1(c)	
2(a)	
2(b)	
3(a)	
3(b)	
4(a)	
4(b)	
5(a)	
5(b)	
5(c)	
5(d)	
5(e)	
Total	

Instructions:

- **Fill in your name and roll number on each of the pages.**
- **This exam is for a total of 30 MARKS.**
- **You may use any result covered in class directly without proving it.**
- **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 $\mathscr{F}$.

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

1. Let $A \in \mathscr{F}$ be an event such that $0 < \mathbb{P}(A) < 1$.

   (a) **(2 Marks)**
   Show that for any $B \in \mathscr{F}$,

$$\mathbb{P}(A) \cdot \Big|\mathbb{P}(B|A) - \mathbb{P}(B)\Big| = \mathbb{P}(A^c) \cdot \Big|\mathbb{P}(B|A^c) - \mathbb{P}(B)\Big|.$$

   <u>Solution</u>(i). Use $\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$. We show that RHS=LHS.

$$\begin{aligned}
\mathbb{P}(A^c) \cdot \Big|\mathbb{P}(B|A^c) - \mathbb{P}(B \cap A) - \mathbb{P}(B \cap A^c)\Big| &= \mathbb{P}(A^c) \cdot \Big|\mathbb{P}(B|A^c)\left(1 - \mathbb{P}(A^c)\right) - \mathbb{P}(B \cap A)\Big| \\
&= \mathbb{P}(A^c) \cdot \Big|\mathbb{P}(B|A^c)\mathbb{P}(A) - \mathbb{P}(B|A)\mathbb{P}(A)\Big| \\
&= \mathbb{P}(A^c)\mathbb{P}(A)\Big|\mathbb{P}(B|A^c) - \mathbb{P}(B|A)\Big| \qquad (1) \\
&= \mathbb{P}(A)\Big|\mathbb{P}(B|A^c)\mathbb{P}(A^c) - \mathbb{P}(B|A)\mathbb{P}(A^c)\Big| \\
&= \mathbb{P}(A)\Big|\mathbb{P}(B \cap A^c) - \mathbb{P}(B|A) + \mathbb{P}(B|A)\mathbb{P}(A)\Big| \\
&= \mathbb{P}(A)\Big|\mathbb{P}(B \cap A^c) - \mathbb{P}(B|A) + \mathbb{P}(B \cap A)\Big| \\
&= \mathbb{P}(A)\Big|\mathbb{P}(B) - \mathbb{P}(B|A)\Big|.
\end{aligned}$$

   <u>Solution</u>(ii). We first write LHS as $\Big|\mathbb{P}(B \cap A) - \mathbb{P}(A)\mathbb{P}(B)\Big|$ and RHS as $\Big|\mathbb{P}(B \cap A^c) - \mathbb{P}(A^c)\mathbb{P}(B)\Big|$. We now show that RHS=LHS, where the idea is to simplify the terms involving complements, as LHS doesn't have complements.

$$\begin{aligned}
\Big|\mathbb{P}(B \cap A^c) - \mathbb{P}(A^c)\mathbb{P}(B)\Big| &= \Big|\mathbb{P}(B \cap A^c) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B)\Big| \\
&= \Big|\mathbb{P}(B)\left(\mathbb{P}(A^c|B) - 1\right) + \mathbb{P}(A)\mathbb{P}(B)\Big| \\
&= \Big| - \mathbb{P}(B)\mathbb{P}(A|B) + \mathbb{P}(A)\mathbb{P}(B)\Big| \\
&= \Big| - \mathbb{P}(A \cap B) + \mathbb{P}(A)\mathbb{P}(B)\Big|.
\end{aligned}$$

   (b) **(2 Marks)**
   Show that for any $B \in \mathscr{F}$,

$$0 \leq \Big|\mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B)\Big| \leq \frac{1}{4}.$$

Solution(i).

$$\begin{aligned}
\left|\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)\right| &= \mathbb{P}(A)\left|\mathbb{P}(B|A)-\mathbb{P}(A)\mathbb{P}(B)\right| \\
&= \mathbb{P}(A^c)\mathbb{P}(A)\left|\mathbb{P}(B|A^c)-\mathbb{P}(B|A)\right| \text{ (From (1))} \\
&\le \mathbb{P}(A^c)\mathbb{P}(A) \ (\because \left|\mathbb{P}(B|A)-\mathbb{P}(B|A^c)\right| \le 1) \\
&= (1-\mathbb{P}(A))\,\mathbb{P}(A) \\
&\overset{(1)}{\le} 1/4.
\end{aligned}$$

Remark1: To arrive at the inequality (1), we can either solve $\max_{p\in(0,1)} p(1-p)$ by either setting the gradients to 0 (allowed when either it's an unconstrained optimization or when the feasibility set is an open set, here it is the latter) or draw this quadratic to check its maximum value or use the AM-GM inequality that holds for non-negative numbers. A special result from the AM-GM inequality is that when the sum of non-negative numbers is a constant (1 in this case), their product is maximized when each of those no.s are equal (1/2 in this case).

Solution(ii). We separately show the bounds $\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)\le 1/4$ and $-1/4\le \mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)$.

$$\begin{aligned}
\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B) &\overset{(1)}{\le} \mathbb{P}(A\cap B)\,(1-\mathbb{P}(A\cap B)) \\
&\overset{(2)}{\le} 1/4,
\end{aligned}$$

where in (1), we have used that by monotonicity, $-\mathbb{P}(A)\mathbb{P}(B)\le -\mathbb{P}(A\cap B)^2$ and (2) follows from Remark1.

For showing the other bound, we can not relate the event of intersection and the individual events via monotonicity. So we start with a general relationship that $A=(A\setminus B)\cup(A\cap B) \implies \mathbb{P}(A\cap B)=\mathbb{P}(A)-\mathbb{P}(A\cap B^c)$. Using this, we get

$$\begin{aligned}
\mathbb{P}(A)\mathbb{P}(B)-\mathbb{P}(A\cap B) &= \mathbb{P}(A)\mathbb{P}(B)+\mathbb{P}(A\cap B^c)-\mathbb{P}(A) \\
&= -\mathbb{P}(A)\mathbb{P}(B^c)+\mathbb{P}(A\cap B^c) \\
&\le 1/4 \text{ (From the first bound proved above with } B \text{ replaced with } B^c).
\end{aligned}$$

(c) **(2 Marks)**
For any $B\in\mathscr{F}$, show that the probability that **exactly** one of the events A or B occurs is given by

$$\mathbb{P}(A)+\mathbb{P}(B)-2\,\mathbb{P}(A\cap B).$$

Solution. We need $\mathbb{P}((A\cup B)\setminus(A\cap B))=\mathbb{P}(A\cup B)-\mathbb{P}(A\cap B)$ where we used that $A\cap B\subseteq A\cup B$. Now, using the inclusion exclusion formula for $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)$, we get the result.

2. (a) **(3 Marks)**

Let $X_1, X_2, \dots \overset{\text{i.i.d.}}{\sim}$ Bernoulli$(1/2)$.
Let X denote the sum

$$X = \sum_{n=1}^{\infty} \frac{X_n}{2^n}.$$

Show that $X \sim$ Unif$((0,1))$.

Hint:

Fix $x \in (0,1)$.
Let $0.x_1x_2x_3\dots$ denote the infinite binary expansion of x, where $x_n \in \{0,1\}$ for each $n \in \mathbb{N}$.
Then, x may be expressed as

$$x = \sum_{n=1}^{\infty} \frac{x_n}{2^n}.$$

Furthermore,

$$\begin{aligned}\mathbb{P}(\{X \le x\}) &= \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\}\right) \\ &= \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \bigcap \left\{X_1 < x_1\right\}\right) + \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \bigcap \left\{X_1 = x_1\right\}\right).\end{aligned}$$

Simplify each of the probability terms in the second line above, and proceed recursively.

Solution. Showing $X \sim$ Unif$((0,1))$ is same as showing $F_X(x) = x$ whenever $x \in (0,1)$, where we have used that $X \in (0,1)$.
We first expand on the hint given (not part of the proof asked). Firstly, as $x \in (0,1)$, it can be expressed as the binary expansion to match the form in which X is written. Then, as there are multiple random variables involved, the idea is to condition on them and simplify. As we can see a correspondence between X_i and x_i, the hint first considers the intersection of the desired event with the disjoint events involving X_1 and x_1: $\{X_1 < x_1\}$, $\{X_1 = x_1\}$ and $\{X_1 > x_1\}$.
We can write, $\mathbb{P}(\{X \le x\}) = \mathbb{P}(\{X \le x\} \cap \{X_1 < x_1\}) + \mathbb{P}(\{X \le x\} \cap \{X_1 = x_1\}) + \mathbb{P}(\{X \le x\} \cap \{X_1 > x_1\})$.
The last term is, however, zero (and hence omitted from the hint). This is because $X_1 > x_1$ can happen only when $X_1 = 1$ and $x_1 = 0$ as both $X_1, x_1 \in \{0,1\}$. Now, $X_1 = 1 \implies X \ge 1/2$ and $x_1 = 0 \implies x \le 1/2$, which in turn implies that $X \ge x$. Thus, $\mathbb{P}(\{X \le x\} \cap \{X_1 > x_1\}) \le \mathbb{P}(\{X \le x\} \cap \{X \ge x\}) = \mathbb{P}(\{X = x\}) = 0$ as the first bits $X_1 \ne x_1$. (To see why $x_1 = 0 \implies x \le 1/2$, note that $x = 0.0x_2 \dots \le \sum_{i=2}^{\infty} \frac{1}{2^i} = \frac{1}{2}$.)

We now start the proof asked.

$$\begin{aligned}&\mathbb{P}(\{X \le x\}) \\ &= \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \bigcap \{X_1 < x_1\}\right) + \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \cap \{X_1 = x_1\}\right) \\ &= \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_1 < x_1\}\right) \mathbb{P}(\{X_1 < x_1\}) + \mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_1 = x_1\}\right) \mathbb{P}(\{X_1 = x_1\})\end{aligned}$$

Now, we analyze each of the terms separately.

- As $X_1 \sim$ Ber$(1/2)$ and $x_1 \in \{0,1\}$, $\mathbb{P}(\{X_1 = x_1\}) = 1/2$.
- $\mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_1 = x_1\}\right) = \mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\}\right)$ as $\frac{X_1}{2}$ in the LHS cancels with $\frac{x_1}{2}$ in the RHS.

- $X_1 < x_1 \begin{cases} \text{happens w.p. } 1/2 \text{ if } x_1 = 1 \\ \text{happens w.p. } 0 \text{ if } x_1 = 0. \end{cases}$ Thus $\mathbb{P}(X_1 < x_1) = \frac{x_1}{2}$. We note that writing this expression combining the two cases takes us closer to the final expression desired i.e. $x = \frac{x_1}{2} + \frac{x_2}{2^2} + \ldots$.
- $\mathbb{P}\left(\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} \le \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_1 < x_1\}\right) = 1$. This is because the event $X_1 < x_1$ happens when $X_1 = 0$ and $x_1 = 1$. With this, the LHS of the event, $\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{1}{2^n} \le \frac{1}{2}$ and the RHS of the event, $\frac{1}{2} + \sum_{n=2}^{\infty} \frac{x_n}{2^n} \ge \frac{1}{2}$, so LHS$\le$RHS with probability 1.

With these, we get that $\mathbb{P}(\{X \le x\}) = \frac{x_1}{2} + \frac{1}{2}\mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\}\right)$. We now repeat the procedure of taking intersection with events involving X_2 and x_2.

$$\begin{aligned}
&\mathbb{P}(\{X \le x\}) \\
&= \frac{x_1}{2} + \frac{1}{2}\mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\}\right) \\
&= \frac{x_1}{2} + \frac{1}{2}\left(\mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\} \bigcap \{X_2 < x_2\}\right) + \mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\} \cap \{X_2 = x_2\}\right)\right) \\
&= \frac{x_1}{2} + \frac{1}{2}\left(\mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_2 < x_2\}\right) \mathbb{P}(\{X_2 < x_2\})\right. \\
&\qquad\qquad \left. + \mathbb{P}\left(\left\{\sum_{n=2}^{\infty} \frac{X_n}{2^n} \le \sum_{n=2}^{\infty} \frac{x_n}{2^n}\right\} \mid \{X_2 = x_2\}\right)\right)\mathbb{P}(\{X_2 = x_2\}) \\
&= \frac{x_1}{2} + \frac{1}{2}\left(\frac{x_2}{2} + \frac{1}{2}\mathbb{P}\left(\left\{\sum_{n=3}^{\infty} \frac{X_n}{2^n} \le \sum_{n=3}^{\infty} \frac{x_n}{2^n}\right\}\right)\right) \\
&= \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{1}{2^2}\mathbb{P}\left(\left\{\sum_{n=3}^{\infty} \frac{X_n}{2^n} \le \sum_{n=3}^{\infty} \frac{x_n}{2^n}\right\}\right)
\end{aligned}$$

Proceeding recursively, we get $\mathbb{P}(\{X \le x\}) = \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3}{2^3} + \cdots = x$. This matches the CDF of a Unif((0, 1)) distribution.

(b) **(3 Marks)**

Fix $q \in (0,1)$. Let $U \sim$ Unif$((0,1))$, and let

$$X = \lfloor \log_q U \rfloor + 1,$$

where $\lfloor x \rfloor$ denotes the largest integer lesser than or equal to x (for e.g., $\lfloor 0.3 \rfloor = 0$, $\lfloor 4.99 \rfloor = 4$, $\lfloor 2 \rfloor = 2$, and so on).
Here, $\log_q U$ denotes the logarithm of U to the base q.
Determine the PMF of X.

Hint:

List down the possible values of $\lfloor \log_q U \rfloor$.

Solution. We can see that X takes only integer values. Further, for any $u \sim$ Unif$((0,1))$, $\log_q u = \frac{\log_e u}{\log_e q}$. As $u < e$ and $q < e$, both the numerator and denominator are negative, making X a positive integer valued random variable (as $u < 1$ always). As we know the distribution of U so we want to relate X in terms of U.
For $k > 0$, $X = k \iff \lfloor \log_q U \rfloor = k-1 \iff k-1 \le \log_q U < k \iff k-1 \le \frac{\log_e u}{\log_e q} < k \iff k \log_e q < \log_e u \le (k-1)\log_e q$ ($\because \log_e q < 0$) $\iff e^{k \log_e q} < u \le e^{(k-1)\log_e q}$ ($\because \log_e(\cdot)$ is an increasing function).

$$\begin{aligned}
\mathbb{P}(\{X = k\}) &= \mathbb{P}(\{q^k < U \le q^{k-1}\}) \\
&= \mathbb{P}(\{q^k \le U \le q^{k-1}\}) \because (U \text{ is a continuous R.V.}) \\
&= F_U(q^{k-1}) - F_U(q^k) \\
&= q^{k-1}(1-q).
\end{aligned}$$

This also shows that X is a Geometric random variable with parameter $1-q$. We note that this is the way of simulating a Geometric R.V. (takes countably infinite no. of values) using the generating function $\lfloor \log_q u \rfloor$. Starting with the generating function $\lfloor \log_q (1-u) \rfloor$ will simulate the Geometric random variable with parameter q.

Note that instead of converting in terms of $\log_e$, we can also solve by taking into account that $\log_q u$ is a decreasing function of u for $u \in (0,1)$ and $q \in (0,1)$.

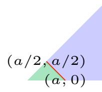
Name:
Roll Number:
Department:
Program: BTech / MTech TA / MTech RA / PhD (Tick one)

3. Let X, Y be jointly continuous with the joint PDF

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

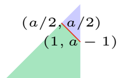
(a) **(3 Marks)**
Determine the PDF of $Z = X + Y$.

Solution. We start with the deriving the CDF of Z at a. We first note that $F_Z(a) = 1$ when $a \ge 2$ and $F_Z(a) = 0$ when $a \le 0$. We thus consider $a \in (0, 2)$ in the following derivation.
To evaluate $F_Z(a) = \mathbb{P}(\{X + Y \le a\})$, we need to integrate the joint PDF over $\{(x, y) : x + y \le a\}$. The PDF takes non-zero values over the triangle with coordinates (0, 0), (1, 0) and (1, 1). Hence, we need to integrate the PDF over the intersection of this triangular region and the half-space obtained with the line $y \le a - x$. This line $y = a - x$ cuts the x-axis at $(a, 0)$ so it is clear that the intersection region will change depending on whether $a \in (0, 1)$ or $a \in (1, 2)$. Hence, we separately analyze the two cases.

Case1: $a \in (0, 1)$. The region of intersection has been shown in green: 

.

$$\begin{aligned} F_Z(a) &= \int_0^{a/2} \int_0^x 1/x \, \mathrm{d}y \, \mathrm{d}x + \int_{a/2}^{a} \int_0^{a-x} 1/x \, \mathrm{d}y \, \mathrm{d}x \\ &= \int_0^{a/2} \mathrm{d}x + \int_{a/2}^{a} \frac{a - x}{x} \, \mathrm{d}x \\ &= a/2 + a(\ln(a) - \ln(a/2)) - a/2 = a \ln(2). \end{aligned}$$

Case2: $a \in (1, 2)$. The region of intersection has been shown in green: 

$$\begin{aligned} F_Z(a) &= \int_0^{a/2} \int_0^x 1/x \, \mathrm{d}y \, \mathrm{d}x + \int_{a/2}^{1} \int_0^{a-x} 1/x \, \mathrm{d}y \, \mathrm{d}x \\ &= a/2 + a(-\ln(a/2)) - 1 + a/2 = a - 1 - a \ln(a/2). \end{aligned}$$

Now, the PDF of Z can be derived by differentiating wrt a. $f_Z(a) = \begin{cases} \ln(2) & \text{if } a \in (0,1); \\ \ln(2/a) & \text{if } a \in (1,2); \\ 0 & a \notin (0,2). \end{cases}$

(We could have also used Leibnitz rule for differentiation under integration to directly obtain the PDF.)

(b) **(1 Mark)**
Compute $\mathbb{P}(\{Z \le 1/2\})$.

Solution. We evaluate the CDF derived for $a \in (0, 1)$ at $a = (1/2)$ to get $\frac{1}{2} \ln(2)$ as the answer.

4. Numbers from $[0,1]$ are picked uniformly, independently, and sequentially over time.
Let X_n denote the number picked at time n, where $n \in \{0,1,2,\ldots\}$.
Let N be the random variable defined as

$$N = \min\{n \geq 1 : X_n < X_0\}.$$

That is, N denotes the first time index $n \geq 1$ at which the value of X_n goes below the value of X_0.

(a) **(3 Marks)**
For any fixed $n \in \mathbb{N}$, determine $\mathbb{P}(\{N = n\})$.
Hint:
The event that $N = n$ is identical to the event that

$$X_1 \geq X_0 \quad \textit{and} \quad X_2 \geq X_0 \quad \textit{and} \quad \cdots \quad \textit{and} \quad X_{n-1} \geq X_0 \quad \textit{and} \quad X_n < X_0.$$

Solution. Let $S(x_0)$ denote the region $\{(x_1, x_2, \ldots, x_n) : x_1 \geq x_0,\ x_2 \geq x_1,\ \ldots, x_{n-1} \geq x_0, x_n < x_0\}$.

$$\begin{aligned}
\mathbb{P}(\{N = n\}) &= \int_0^1 \int_{x_1,\cdots,x_n \in S(x_0)} f_{X_0,\ldots,X_n}(x_0,\cdots,x_n)\,\mathrm{d}x_0\,\mathrm{d}x_1\,\ldots\mathrm{d}x_n \\
&= \int_0^1 f_{X_0}(x_0) \int_{x_1,\cdots,x_n \in S(x_0)} f_{X_1,\ldots,X_n|X_0=x_0}(x_1,\cdots,x_n)\,\mathrm{d}x_1\,\cdots\mathrm{d}x_n\,\mathrm{d}x_0 \\
&= \int_0^1 f_{X_0}(x_0)\,\mathbb{P}(\{X_1 \geq X_0,\cdots,X_{n-1} \geq X_0, X_n < X_0\}|\{X_0 = x_0\})\,\mathrm{d}x_0 \\
&= \int_0^1 f_{X_0}(x_0)\,\mathbb{P}(\{X_1 \geq x_0,\cdots,X_{n-1} \geq x_0, X_n < x_0\})\,\mathrm{d}x_0 \\
&= \int_0^1 f_{X_0}(x_0)\,\Pi_{i=1}^{n-1}\,\mathbb{P}(\{X_i \geq x_0\})\,\mathbb{P}(\{X_n < x_0\})\,\mathrm{d}x_0 \text{ (Using independence of } X_i's,\ i \in \mathbb{N}) \\
&= \int_0^1 (1-x_0)^{n-1}\,x_0\,\mathrm{d}x_0 \ (\because X_0,\cdots X_n \text{ sampled uniformly from } [0,1]) \\
&= \int_0^1 -(1-x_0)^{n-1}\,(1-x_0-1)\,\mathrm{d}x_0 \\
&= \int_0^1 -(1-x_0)^{n} + (1-x_0)^{n-1}\,\mathrm{d}x_0
\end{aligned}$$

To solve the integration without using any special formula, we can do a change of variable. Let $x := 1 - x_0 \implies \mathrm{d}x = -\mathrm{d}x_0$ and when $x_0 = 0, x = 1$; when $x_0 = 1, x = 0$.

$$\mathbb{P}(\{N = n\}) = \int_1^0 x^n - x^{n-1}\,\mathrm{d}x = -\frac{1}{n+1} + \frac{1}{n} = \frac{1}{n(n+1)}$$

(b) **(1 Mark)**
Compute $\mathbb{P}(\{N > 2\})$.

Solution. $\mathbb{P}(\{N > 2\}) = 1 - \mathbb{P}(\{N = 1\}) - \mathbb{P}(\{N = 2\}) = 1 - 1/2 - 1/6 = 1/3.$

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

5. Let $X_1, X_2 \overset{\text{i.i.d.}}{\sim}$ Exponential(λ).

Two Ph.D. students of IIT Hyderabad (let us call them S_1 and S_2) are on a mission to come up with their own definitions for what it means to "condition on" the event $\{X_1 = X_2\}$.

- **First Definition:** Student S_1 reasons that $X_1 = X_2$ if and only if $\frac{X_1}{X_2} = 1$, and therefore finds it apt to define conditioning on the event $\{X_1 = X_2\}$ as conditioning on the event $\left\{\frac{X_1}{X_2} = 1\right\}$.
- **Second Definition:** Student S_2 reasons that $X_1 = X_2$ if and only if $X_1 - X_2 = 0$, and therefore finds it apt to define conditioning on the event $\{X_1 = X_2\}$ as conditioning on the event $\left\{X_1 - X_2 = 0\right\}$.

(a) **(2 Marks)**
Show that $\{X_1 = X_2\} = \{\omega \in \Omega : X_1(\omega) = X_2(\omega)\} \in \mathscr{F}$.
Furthermore, argue that $\mathbb{P}(\{X_1 = X_2\}) = 0$.

Solution. As $X_1,\ X_2$ are random variables, $X_1 - X_2$ is also a random variable as it is a continuous function of two random variables. To prove the desired result, we simply use that for the singleton Borel set $\{0\}$, the preimage of $X_1 - X_2$ should belong to $\mathscr{F}$ i.e. $\{\omega \in \Omega : X_1(\omega) - X_2(\omega) = 0\} \in \mathscr{F}$.

Further, we know that $\mathbb{P}(\{X_1 - X_2\}) := \mathbb{P}_{X_1 - X_2}(\{0\}) = 0$ where the last equality uses that $X_1 - X_2$ is a continuous random variable (as $X_1,\ X_2$ are independently continuous random variables).

(b) **(2 Marks)**
Determine the joint PDF of $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_2}$.

Solution. (Same question as in Assignment 6, Q5.) The transformation function $g(x_1, x_2) := \begin{bmatrix} x_1 + x_2 \\ x_1/x_2 \end{bmatrix}$, where $g_1(x_1, x_2) = x_1 + x_2,\ g_2 = x_1/x_2$ are differentiable with continuous first order partial derivatives.
We now test if g is one-one. Let $g(x_1, x_2) = g(x_1', x_2')$. On solving the equations $x_1 + x_2 = x_1' + x_2'$ and $x_1/x_2 = x_1'/x_2'$, we get $x_1 = x_1'$ and $x_2 = x_2'$. This implies that g is a one-one function.
The Jacobian matrix for g is $J_g(x_1, x_2) = \begin{bmatrix} 1 & 1 \\ 1/x_2 & -x_1/x_2^2 \end{bmatrix}$. The corresponding Jacobian = Det$(J_g(x_1, x_2)|) = -\frac{x_1 + x_2}{x_2^2}$.
As $X_1, X_2 \sim$ Exp(λ), the Jacobian is never 0.
Now, we can apply the transformation formula. To find $g^{-1}(y_1, y_2)$, we solve $x_1 + x_2 = y_1$ and $x_1/x_2 = y_2$ for x_1 and x_2. Thus, $g^{-1}(y_1, y_2) = \begin{bmatrix} \frac{y_1 y_2}{1+y_2} \\ \frac{y_1}{1+y_2} \end{bmatrix}$ and Det $\left(J_g\left(\frac{y_1 y_2}{1+y_2}, \frac{y_1}{1+y_2}\right)\right) = -\frac{(1+y_2)^2}{y_1}$. Finally, we put these in the transformation formula, $f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{f_{X_1,X_2}(g^{-1}(y_1,y_2))}{|\text{Det}(J_g(g^{-1}(y_1,y_2)))|}, & y_1 > 0, y_2 \geq 0 \\ 0 & \text{Otherwise} \end{cases}$ and use that X_1 and X_2 are independent, we get $f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{y_1 e^{-\lambda y_1}}{(y_2+1)^2}\lambda^2, & y_1 > 0, y_2 \geq 0 \\ 0 & \text{Otherwise} \end{cases}$.

(c) **(2 Marks)**
Determine the conditional PDF of Y_1, conditioned on the event $\{Y_2 = 1\}$.

Solution. To find $f_{Y_1|Y_2=y}$, we need to first find $f_{Y_2=y}$ by marginalizing out Y_1 in the joint PDF f_{Y_1,Y_2}.
We get

$$f_{Y_2}(y_2) = \begin{cases} \int_0^\infty \frac{y_1 e^{-\lambda y_1}}{(y_2+1)^2}\lambda^2 \, \mathrm{d}y_1, & y_2 > 0 \\ 0, & \text{otherwise} \end{cases} \overset{(1)}{=} \begin{cases} \frac{1}{(1+y_2)^2}, & y_2 > 0 \\ 0, & \text{otherwise,} \end{cases}$$

where step (1) uses integration by parts rule, taking the first function as $f(y) = y$ and the second function as $f(y) = e^{-\lambda y}$.

The conditional PDF, $f_{Y_1|Y_2=y_2} = \begin{cases} \lambda^2 y_1 e^{-\lambda y_1}, & y_1 > 0 \\ 0, & \text{Otherwise,} \end{cases}$ which is just the Erlang distribution. The conditional PDF is the same for all conditioned events $\{Y = y_2\}$ when $y_2 \geq 0$.

(d) **(2 Marks)**
Determine the joint PDF of Y_1 (as defined above) and $Y_3 = X_1 - X_2$.

Solution. The transformation function $g(x_1, x_2) := \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}$, where $g_1(x_1, x_2) = x_1 + x_2,\ g_2 = x_1 - x_2$ are differentiable with continuous first order partial derivatives. It is easy to see that g is one-one following the method in part (c). The Jacobian matrix for g is $J_g(x_1, x_2) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. The corresponding Jacobian is $\text{Det}(J_g(x_1, x_2)) = -2 \neq 0$.
Now, we can apply the transformation formula. To find $g^{-1}(y_1, y_3)$, we solve $x_1 + x_2 = y_1$ and $x_1 - x_3 = y_3$ for x_1 and x_2. Thus, $g^{-1}(y_1, y_3) = \begin{bmatrix} \frac{1}{2}(y_1 + y_3) \\ \frac{1}{2}(y_1 - y_3) \end{bmatrix}$. Finally, we put these in the transformation formula, $f_{Y_1,Y_3}(y_1, y_3) = \begin{cases} \frac{f_{X_1,X_2}(g^{-1}(y_1,y_3))}{|\text{Det}(J_g(g^{-1}(y_1,y_3)))|}, & y_1 + y_3 \geq 0;\ y_1 - y_3 \geq 0 \\ 0 & \text{Otherwise} \end{cases}$ and use that X_1 and X_2 are independent, we get
$f_{Y_1,Y_3}(y_1, y_3) = \begin{cases} \frac{\lambda^2 e^{-\lambda y_1}}{2} & y_1 \geq |y_3| \\ 0 & \text{Otherwise} \end{cases}$.

(e) **(2 Marks)**
Determine the conditional PDF of Y_1, conditioned on the event $\{Y_3 = 0\}$.

Solution. To find $f_{Y_1|Y_3=y_3}$, we need to first find $f_{Y_3=y_3}$ by marginalizing out Y_1 in the joint PDF f_{Y_1,Y_3}.
We get $f_{Y_3}(y_3) = \int_{|y_3|}^{\infty} \frac{\lambda^2 e^{-\lambda y_1}}{2} = \frac{\lambda e^{-\lambda|y_3|}}{2}$. The conditional PDF $f_{Y_1|Y_3=y_3} = \begin{cases} \lambda e^{-\lambda(y_1-|y_3|)} & y_1 \geq |y_3| \\ 0 & \text{Otherwise} \end{cases}$.
The conditional PDF $f_{Y_1|Y_3=0} = \begin{cases} \lambda e^{-\lambda y_1} & y_1 \geq 0 \\ 0 & \text{Otherwise} \end{cases}$.
The answers in parts (c) and (e) are different. This happens when conditioning on a 0 probability event. $\{Y_3 = 0\} \implies X_2 - \epsilon \leq X_1 \leq X_2 + \epsilon$ for $\epsilon \to 0$, and $\{Y_2 = 1\} \implies (1-\epsilon)X_2 \leq X_1 \leq (1+\epsilon)X_2$ for $\epsilon \to 0$, which are different inequalities.

Finally, show that the conditional PDFs in parts (c) and (e) are different.

Here, conditioning according to student S_1's definition leads to a different answer than according to student S_2's definition. The above problem shows that when conditioning on a zero probability event, one must exercise care to specify the exact definition of conditioning.