Probability and Random Processes — Monsoon 2023

Assignment 5 Solutions

PRP TAs

November 17, 2023

Problem 1

(a)

P(A), or P(X $\geq$ 2), is given by probability law on $f_X(x)$

$$\begin{aligned} P(X \geq 2) &= \int_2^{\infty} f_X(x).dx \\ &= \int_2^3 \frac{x}{4}.dx \\ &= \frac{1}{8}x^2 \Big|_2^3 \\ \implies P(X \geq 2) &= \frac{5}{8} \end{aligned}$$

(b)

By using definition of conditional expectation for joint pdf, conditioned on discrete random variable.

$$\begin{aligned} f_{X|A}(x) &= \frac{f_X(x)}{P(A)} \\ f_{X|A}(x) &= \frac{\frac{x}{4}}{\frac{5}{8}} \\ &= \frac{2x}{5}(1 \leq x \leq 3) \end{aligned}$$

(c)

$$\begin{aligned}
E[X|A] &= \int x.f_{X|A}(x).dx \\
&= \int_2^3 X.\frac{2x}{5}.dx \\
&= \int_2^3 \frac{2x^2}{5}.dx \\
&= \left.\frac{2x^3}{15}\right|_2^3 \\
\implies E[X|A] &= \frac{38}{15}
\end{aligned}$$

(d)

$$\begin{aligned}
E[X] &= \int x.f_X(x).dx \\
&= \int_1^3 x.\frac{x}{4}.dx \\
&= \int_1^3 \frac{x^2}{4}.dx \\
\implies E[X] &= \frac{13}{6}
\end{aligned}$$

Problem 2

We have two continuous random variables with joint pdf given as follows,

$$f_{X,Y}(x,y) = \begin{cases} 4x^2 & 0 < y < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

We find the marginal distribution of Y, $f_Y(y)$, as follows,

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \\
&= \int_y^1 4x^2\,dx \\
&= \left.\frac{4}{3}x^3\right|_y^1 \\
&= \frac{4}{3}(1-y^3)
\end{aligned}$$

Hence, $f_Y(y) = \frac{4}{3}(1 - y^3)$.
Then, we can compute the conditional distribution $f_{X|Y}(x|y)$,

$$
\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \begin{cases} \frac{3x^2}{1-y^3} & 0 < y < x < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Problem 3

We utilize the fact that the pdf of X is two-sided exponential. We can use total expectation theorem to get expectation of X in terms of expectation of individual exponential RVs. Using total expectation theorem, we can write that

$$E[X] = E[X|X \geq 0]P(X \geq 0) + E[X|X < 0]P(X < 0)$$

Firstly, we calculate the conditional pdfs of X, for which we first calculate the probabilities:

$$
\begin{aligned}
P(X \geq 0) &= \int_0^\infty f_X(x)\,dx = p \\
P(X < 0) &= \int_{-\infty}^0 f_X(x)\,dx = 1 - p
\end{aligned}
$$

Thus the Conditional pdfs are:

$$
\begin{aligned}
f_{X|X\geq 0}(x) &= \lambda e^{-\lambda x} \text{ for } x \geq 0 \\
f_{X|X<0}(x) &= \lambda e^{\lambda x} \text{ for } x < 0
\end{aligned}
$$

These two conditional pdfs are of exponential and negative exponential RV respectively. We know that the expectation of exponential RV is the inverse of parameter λ ,i.e., $\frac{1}{\lambda}$ and the variance is $\frac{1}{\lambda^2}$. Using the expression $Var(X) = E[X^2] - (E[X])^2$, we can see that the expectation of the square of exponential RV is $\frac{2}{\lambda^2}$. Using these facts, we write the conditional expectation of X and X^2 as:

$$
\begin{aligned}
E[X|X \geq 0] &= \frac{1}{\lambda} \\
E[X|X < 0] &= \frac{-1}{\lambda} \\
E[X^2|X \geq 0] &= \frac{2}{\lambda^2} \\
E[X^2|X < 0] &= \frac{2}{\lambda^2}
\end{aligned}
$$

Thus, the expectation of X can be calculated as:

$$
\begin{aligned}
E[X] &= E[X|X \geq 0]P(X \geq 0) + E[X|X < 0]P(X < 0) \\
&= p(\frac{1}{\lambda}) + (1-p)(\frac{-1}{\lambda}) \\
&= \frac{2p-1}{\lambda}
\end{aligned}
$$

Similarly, the expectation of X^2 comes out to be:

$$\begin{aligned} E[X^2] &= E[X^2|X \geq 0]P(X \geq 0) + E[X^2|X < 0]P(X < 0) \\ &= p(\frac{2}{\lambda^2}) + (1-p)(\frac{2}{\lambda^2}) \\ &= \frac{2}{\lambda^2} \end{aligned}$$

Thus, the variance of X is:

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \\ &= \frac{2}{\lambda^2} - \frac{(2p-1)^2}{\lambda^2} \\ &= \frac{4p - 4p^2 + 1}{\lambda^2} \end{aligned}$$

Problem 4

(a), (b), (c)

P(X_n is record to date)

$$= P(X_n > \max\{X_1, \dots X_{n-1}\})$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_n} f_{X_1,\dots X_n}(x_1,x_2\dots x_n)\,dx_1\,dx_2\dots dx_n$$

Heres X_n's are independent

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_n} f_{X_1}(x_1)\,f_{X_2}(x_2)\dots f_{X_n}(x_n)\,dx_1\,dx_2\dots dx_n$$

Here X_n's are identically distributed

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_n} f_X(x_1)\,f_X(x_2)\dots f_X(x_n)\,dx_1\,dx_2\dots dx_n$$

$$= \int_{-\infty}^{\infty} f_X(x_n)\left(\int_{-\infty}^{x_n} f_X(x_{n-1})\,dx_{n-1}\right)\dots\left(\int_{-\infty}^{x_n} f_X(x_1)\,dx_1\right)dx_n$$

By definition of CDF

$$= \int_{-\infty}^{\infty} f_X(x_n)\,F_X(x_n)\dots F_X(x_n)\,dx_n$$

$$= \int_{-\infty}^{\infty} f_{X_n}(x_n)\left(F_X(x_n)\right)^{n-1}dx_n$$

↳ Calculate By Considering $F_X(x_n) = t$, $f_X(x_n)\,dx = dt$.

P(X_n record to date) $= \frac{1}{n}$.

(c) Event A_n is record to date.

A_{n-1}: X_{n-1} is record to date.

A_n & A_{n-1} are independent events

A_n & $\overline{A_{n-1}}$ are independent Since ($A \perp\!\!\!\perp B \Rightarrow A \perp\!\!\!\perp \bar{B}$)

By this,

$$P_r(X_m \text{ is 1st record to date}) = P_r(X_2 \text{ is not r-t-d} \cap \dots \cap X_m \text{ is r-t-d})$$

$$= P_r(\bar{A_2} \cap \dots A_m)$$

$$= P_r(\bar{A_2})\,P_r(\bar{A_3})\dots P_r(A_m)$$

$$= \left(1-\tfrac{1}{2}\right)\left(1-\tfrac{1}{3}\right)\dots\left(\tfrac{1}{m}\right)$$

$$= \frac{1}{m(m-1)}$$

$$P_r(N \geq n) = 1 - P_r(N_1 \leq n)$$

$$= 1 - P_r(N_1 = 2) - P_r(N=3)\dots - P_r(N=n)$$

$$= 1 - \sum_{i=2}^{n} \frac{1}{i(i-1)}$$

$$= 1 - \left(1 - \tfrac{1}{2} + \tfrac{1}{2} - \tfrac{1}{3} - \dots \tfrac{1}{n-1} - \tfrac{1}{n}\right)$$

$$P_r(N \geq n) = \frac{1}{n}$$

Problem 5

(a)

We are given a discrete random variable Y, having CDF,

$$F_Y(y) = 1 - \frac{2}{(y+1)(y+2)}, \quad k \geq 0.$$

We first find the PMF of Y, by evaluating the difference equation of its CDF.

$$\begin{aligned}
p_Y(y) &= F_Y(y) - F_Y(y-1) \\
&= \left(1 - \frac{2}{(y+1)(y+2)}\right) - \left(1 - \frac{2}{(y)(y+1)}\right) \\
&= 2\left(\frac{1}{y+1}\right)\left(\frac{1}{y} - \frac{1}{y+2}\right) \\
&= \frac{4}{y(y+1)(y+2)}
\end{aligned}$$

Now we find the expected value of Y, directly from the formula.

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{y=-\infty}^{\infty} y p_Y(y) \\
&= \sum_{y=1}^{\infty} \frac{4}{(y+1)(y+2)} \\
&= 4\sum_{y=1}^{\infty} \left(\frac{1}{y+1} - \frac{1}{y+2}\right) \\
&= 4\Big[\Big(\frac{1}{2} - \cancel{\frac{1}{3}}\Big) + \Big(\cancel{\frac{1}{3}} - \cancel{\frac{1}{4}}\Big) + \dots\Big] \\
&= 2
\end{aligned}$$

Therefore, $\mathbb{E}[Y] = 2$.

(b)

Now we are given a discrete random variable X, which has a conditional PMF,

$$P_{X|Y}(x|y) = \frac{1}{y}, \quad x = 1, \dots, y.$$

We can find the expectation of X by first finding the PMF of X and then using the expectation formula (try it out on your own). However, we can use a more straightforward method, called the law of iterated expectations, illustrated below.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X|Y]] \\
&= \sum_{y=-\infty}^{\infty} \mathbb{E}[X|Y] p_Y(y) \\
&= \sum_{y=-\infty}^{\infty} \Big(\sum_{x=-\infty}^{\infty} x p_{X|Y}(x|y) \Big) p_Y(y) \\
&= \sum_{y=1}^{\infty} \Big(\sum_{x=1}^{y} x \frac{1}{y} \Big) \frac{4}{y(y+1)(y+2)} \\
&= \sum_{y=1}^{\infty} \frac{1}{y} \frac{4}{y(y+1)(y+2)} \frac{y(y+1)}{2} \\
&= \sum_{y=1}^{\infty} \frac{2}{y(y+2)} \\
&= \sum_{y=1}^{\infty} \Big(\frac{1}{y} - \frac{1}{y+2} \Big) \\
&= \Big[\Big(1 - \cancel{\frac{1}{3}}\Big) + \Big(\frac{1}{2} - \cancel{\frac{1}{4}}\Big) + \Big(\cancel{\frac{1}{3}} - \cancel{\frac{1}{5}}\Big) + \dots \Big] \\
&= \frac{3}{2}
\end{aligned}
$$

Therefore, $\mathbb{E}[X] = \frac{3}{2}$.

Problem 6

When both are Continuous RV: We have to prove that the given condition is true, we do it in two parts:
i) Given X and Y are independent, prove $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.

$$
\begin{gathered}
f_{X,Y}(x, y) = f_X(x) f_Y(y) \\
\int_0^{\infty} \int_0^{\infty} f_{X,Y}(x, y)\, dy\, dx = \int_0^{\infty} \int_0^{\infty} f_X(x) f_Y(y)\, dy\, dx \\
F_{X,Y}(x, y) = F_X(x) F_Y(y)
\end{gathered}
$$

Hence, proved.
ii) Given $F_{X,Y}(x, y) = F_X(x)F_Y(y)$, prove that X and Y are independent.

$$\begin{aligned}
F_{X,Y}(x,y) &= F_X(x)F_Y(y) \\
\frac{dF_{X,Y}(x,y)}{dx.dy} &= \frac{d\{F_X(x)F_Y(y)\}}{dx.dy} \\
f_{X,Y}(x,y) &= f_X(x)f_Y(y)
\end{aligned}$$

Hence, proved.
When both are Discrete RV: We have to prove that the given condition is true, we do it in two parts:
i) Given X and Y are independent, prove $F_{X,Y}(x,y) = F_X(x)F_Y(y)$.

$$\begin{aligned}
F_{X,Y}(x,y) &= \sum_x \sum_y P(X = x, Y = y) \\
&= \sum_x \sum_y P(X = x)P(Y = y) \\
&= \sum_x P(X = x) \sum_y P(Y = y) \\
&= F_X(x)F_Y(y)
\end{aligned}$$

Hence, proved.
ii) Given $F_{X,Y}(x,y) = F_X(x)F_Y(y)$, prove that X and Y are independent.

$$\begin{aligned}
P(X = x, Y = y) &= F_{X,Y}(x,y) - F_{X,Y}(x-1,y) - F_{X,Y}(x,y-1) + F_{X,Y}(x-1,y-1) \\
&= F_X(x)F_Y(y) - F_X(x-1)F_Y(y) - F_X(x)F_Y(y-1) + F_X(x-1)F_Y(y-1) \\
&= (F_X(x) - F_X(x-1))(F_Y(y) - F_Y(y-1)) \\
&= P(X = x)P(Y = y)
\end{aligned}$$

Hence, proved.
This proves that two random variables X and Y (either both continuous or both discrete) are independent if and only if $F_{XY}(x,y) = F_X(x)F_Y(y)$, for all x, y.

Problem 7

We will solve this question in two ways:

First let's do it in a more informal but an intuitive way. We will use the fact that $P(X_1 = X_2) = 0$ since X_1, X_2 are continuous random variables. The above fact is true for other pairs of random variable as well i.e. $(X_1, X_3), (X_2, X_3)$.

Let's define $X_iX_jX_k$ to be an ordering on the sampled value of the random variables X_i, X_j, X_k if $x_i < x_j < x_k$. For example if $X_1 = 2, X_2 = 5, X_3 = 4$, then the ordering is $X_1X_3X_2$. Similarly if $X_1 = 3, X_2 = 0, X_3 = 1$, the ordering will be $X_2X_3X_1$. I hope you get the point..

Now, suppose X_1, X_2, X_3 were discrete random variables and the size of sample space of X_1 was n. In that case, you would have to consider n^3 possible triplets of (x_1, x_2, x_3). Now ask yourself the question, in how many cases will the ordering $X_1X_2X_3$ will be followed i.e. $X_i < X_j < X_k$. An intuitive answer will be $\frac{1}{6}$. In fact, probability of orderings

$$P(X_1X_2X_3) = P(X_1X_3X_2) = P(X_2X_1X_3) = P(X_2X_3X_1) = P(X_3X_1X_2) = P(X_3X_2X_1) = \frac{1}{6}$$

This seems quite obvious because X_1, X_2, X_3 are i.i.d. random variables. In fact, its a good exercise to prove the above. Now, returning to the actual question

(a) If A is the tallest i.e. X_1 is the largest. The possible ordering of random variables can be $\_\_X_1$. The blanks can be filled with either X_2 or X_3. Thus, the possible orderings are $X_3X_2X_1$ and $X_2X_3X_1$ and the total probability

$$\text{P(A is tallest)} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

(b) P(A is taller than B | A is taller than C)
$\implies$ X_1 follows X_3 in the ordering. Thus all the possible orderings are $X_3X_1\_$, $X_3\_X_1$ and $\_X_3X_1$. Since there is a single blank which is filled by X_2. As per the question in how many of these X_1 follows X_2 i.e. X_1 comes after a blank. No. of such cases is 2. Thus,

$$\text{P(A is taller than B | A is taller than C)} = \frac{2}{3}$$

(c) P(A is taller than B | B is taller than C)
$\implies$ X_2 follows X_3 and $X1$ follows X_2. Thus only possible orderings are $X_3X_2\_$, $X_3\_X_2$ and $\_X_3X_2$. And the favourable ordering is $X_3X_2X_1$. Thus,

$$\text{P(A is taller than B | B is taller than C)} = \frac{1}{3}$$

Now let's solve it in a more mathematically rigourous way.
Let $f_X(x)$ and $F_X(x)$ be the PDF and the CDF for the X_i random variable.

(a)
$$\begin{aligned}\text{P(A is the tallest Child)} &= P\left(X_1 > X_2, X_3\right)\\ &= \int_{x_1=-\infty}^{\infty}\int_{x_2=-\infty}^{x_1}\int_{x_3=-\infty}^{x_1} f_{X_1X_2X_3}\left(x_1, x_2, x_2\right) dx_3\ dx_2\ dx_1\\ &= \int_{x_1=-\infty}^{\infty}\int_{x_2=-\infty}^{x_1}\int_{x_3=-\infty}^{x_1} f_X\left(x_1\right) f_X\left(x_2\right) f_X\left(x_3\right) dx_3\ dx_2\ dx_1 \ [X_i\text{'s are i.i.d.}]\\ &= \int_{-\infty}^{\infty} F_X\left(x_1\right)^2 \cdot f_X(x_1)\, dx\end{aligned}$$

Let $y = F_X\left(x_1\right) \implies dy = f_x\left(x_1\right) dx$. And $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.

$$\text{P(A is the tallest Child)} = \int_0^1 y^2\ dy = \left[\frac{y^3}{3}\right]_0^1 = \frac{1}{3}$$

(b) P(A is taller than B | A is taller than C) = P($X_1 > X_2 \mid X_1 > X_3$). By Bayes's rule,

$$P(X_1 > X_2|X_1 > X_3) = \frac{P(X_1 > X_2,\ X_1 > X_3)}{P(X_1 > X_3)} = \frac{P(X_1 > X_2, X_3)}{P(X_1 > X_3)}$$

$$P(X_1 > X_2, X_3) = \frac{1}{3} \quad \text{[from part(a)]}$$

$$\begin{aligned}
P\left(X_1 > X_3\right) &= \int_{x_1=-\infty}^{\infty} \int_{x_3=-\infty}^{x_1} f_{X_1 X_3}\left(x_1, x_3\right)\ dx_3\ dx_1 \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_3=-\infty}^{x_1} f_X(x_1) \cdot f_X(x_3)\ dx_3\ dx_1 \\
&= \int_{x_1=-\infty}^{\infty} F_X(x_1) \cdot f_X(x_1)\ dx_1 \\
&= \int_0^1 y \cdot dy \qquad \text{[Using the same substitution as previous part]} \\
&= \left[\frac{y^2}{2}\right]_0^1 \\
&= \frac{1}{2}
\end{aligned}$$

$$\therefore P(X_1 > X_2|X_1 > X_3) = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

(c) P(A is taller than B | B is taller than C) = $P(X_1 > X_2|X_2 > X_3) = \dfrac{P(X_1 > X_2 > X_3)}{P(X_2 > X_3)}$

By similar logic from part (b), $P(X_2 > X_3) = \dfrac{1}{2}$.

$$\begin{aligned}
P\left(X_1 > X_2 > X_3\right) &= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{x_1} \int_{x_3-\infty}^{x_1} f_{X_1 X_2 X_3}\left(x_1, x_2, x_3\right)\ dx_3\ dx_2\ dx_1 \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{x_1} \int_{x_3=-\infty}^{x_2} f_X(x_1)\ f_X(x_2)\ f_X(x_3)\ dx_3\ dx_2\ dx_1 \\
&= \int_{x_1=-\infty}^{\infty} \int_{x_2=-\infty}^{x_1} f_X(x_1)\ f_X(x_2)\ F_X(x_2)\ dx_2\ dx_1 \\
&= \int_{x_1=-\infty}^{\infty} \int_{y=0}^{F_X(x_1)} f_X(x_1)\ y\ dy\ dx_1 \quad \text{[Using } y = F_X(x_2)\text{]} \\
&= \frac{1}{2} \cdot \int_{x_1=-\infty}^{\infty} f_X(x_1) \cdot (F_X(x_1))^2 \quad dx_1 \\
&= \frac{1}{2} \cdot \int_0^1 y^2 \cdot dy \qquad \text{[Using } y = F_X(x_1)\text{]} \\
&= \frac{1}{6}
\end{aligned}$$

$$\therefore P(X_1 > X_2|X_2 > X_3) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$