Queen's University
Department of Mathematics and Statistics

MTHE/STAT 353

Homework 1 Solutions, 2022

- For each question, your solution should start on a fresh page. You can write your solution using one of the following three formats:
  1. (1) Start your solution in the space provided right after the problem statement, and use your own paper if you need extra pages.
  2. (2) Write your whole solution using your own paper, and make sure to number your solution.
  3. (3) Write your solution using document creation software (e.g., Word or LaTeX).
- Write your name and student number on the first page of each solution.
- For each question, photograph or scan each page of your solution (unless your solution has been typed up and is already in electronic format), and combine the separate pages into a single file. Then upload each file (one for each question), into the appropriate box in Crowdmark.

Instructions for submitting your solutions to Crowdmark are also here.

Total Marks : 25

Student Number Name

1. (5 marks) Suppose an urn initially contains one red ball, one blue ball, and one green ball. At each draw, a ball is randomly selected from the urn, replaced, and an additional ball of the same colour as the drawn ball is added to the urn. Thus, after n draws there are $n+3$ balls in the urn. After n draws, let X be the number of times a red ball was drawn, Y the number of times a blue ball was drawn, and Z the number of times a green ball was drawn. Compute the joint probability mass function of the random vector (X, Y, Z). As part of the joint pmf you must also give the support of the distribution!

Solution: Let x, y, z by nonnegative integers summing to n. Then we wish to compute $P(X = x, Y = y, Z = z)$. The probability of a particular sequence of n draws that contains x draws of a red ball, y draws of a blue ball, and z draws of a green ball is

$$\frac{x!y!z!}{3 \times 4 \times \ldots \times (n+2)} = \frac{2x!y!z!}{(n+2)!}$$

To see that the expression on the left above is correct, note that the expression for the probability of any such sequence has denominator $3 \times \ldots \times (n+2)$. The $x!$ accounts for all the numerators of the probabilities for drawing a red ball, regardless of when those draws were. A similar observation for the blue and green balls gives us the $y!$ and $z!$, respectively. Now there are $n!/(x!y!z!)$ such sequences so we end up with

$$P(X = x, Y = y, Z = z) = \frac{n!}{x!y!z!}\frac{2x!y!z!}{(n+2)!} = \frac{2}{(n+1)(n+2)} = \frac{1}{\binom{n+2}{2}}.$$

In other words, the joint distribution of (X, Y, Z) is discrete uniform on the (3-dimensional) support

$$S = \{(x, y, z) : \; x, y, z \text{ are nonnegative integers, and } x + y + z = n\}.$$

The joint pmf of (X, Y, Z) is given by

$$p(x, y, z) = \begin{cases} \binom{n+2}{2}^{-1} & \text{if } (x, y, z) \in S \\ 0 & \text{otherwise.} \end{cases}$$

Student Number Name

2. (5 marks) Let X_1, X_2, X_3 be discrete random variables with joint pmf

$$p_X(x_1,x_2,x_3) = \left(\frac{1}{2}\right)^{x_3}(1-e^{-x_3})^2 e^{-x_3(x_1+x_2-2)},$$

for $x_1, x_2, x_3 = 1, 2, 3, \ldots$ and $p_X(x_1,x_2,x_3) = 0$ otherwise. Find the marginal pmf of X_1.

Solution: Let $x_1 \in \mathbb{N}$ be fixed. Sum over x_2 first:

$$\begin{aligned}\sum_{x_2=1}^{\infty} p_X(x_1,x_2,x_3) &= \left(\frac{1}{2}\right)^{x_3}(1-e^{-x_3})e^{-x_3(x_1-1)}\sum_{x_2=1}^{\infty}(1-e^{-x_3})(e^{-x_3})^{x_2-1} \\ &= \left(\frac{1}{2}\right)^{x_3}(1-e^{-x_3})e^{-x_3(x_1-1)},\end{aligned}$$

since the sum, being the sum over all probabilities of a Geometric$(1-e^{-x_3})$ distribution, is equal to one. Next, sum over x_3:

$$\begin{aligned}\sum_{x_3=1}^{\infty}\sum_{x_2=1}^{\infty} p_X(x_1,x_2,x_3) &= \sum_{x_3=1}^{\infty}\left(\frac{1}{2}\right)^{x_3}(1-e^{-x_3})e^{-x_3(x_1-1)} \\ &= \sum_{x_3=1}^{\infty}\left(\frac{e^{-(x_1-1)}}{2}\right)^{x_3} - \sum_{x_3=1}^{\infty}\left(\frac{e^{-x_1}}{2}\right)^{x_3} \\ &= \frac{1}{1-\frac{e^{-(x_1-1)}}{2}} - 1 - \left(\frac{1}{1-\frac{e^{-x_1}}{2}} - 1\right) \\ &= \frac{2}{2-e^{-(x_1-1)}} - \frac{2}{2-e^{-x_1}}.\end{aligned}$$

So the marginal pmf of X_1 is

$$p_{X_1}(x_1) = \begin{cases} \frac{2}{2-e^{-(x_1-1)}} - \frac{2}{2-e^{-x_1}} & \text{for } x_1 = 1, 2, 3, \ldots \\ 0 & \text{otherwise.}\end{cases}$$

Student Number Name

3. (6 marks) Let X_1, X_2, X_3 be continuous random variables with joint pdf

$$f_X(x_1,x_2,x_3) = \frac{1}{\sqrt{2\pi}}e^{-(x_1-x_3)^2/2}\frac{1}{\sqrt{2\pi}}e^{-(x_2-x_3)^2/2}\frac{1}{\sqrt{2\pi}}e^{-x_3^2/2},$$

for $-\infty < x_1, x_2, x_3 < \infty$. Find the joint marginal pdf of (X_1, X_2) and the marginal pdf of X_1.

Solution: To get the joint marginal pdf of (X_1, X_2) we need to integrate out x_3. Combining the exponents above we get a quadratic function of x_3, which we will write as $-\frac{1}{2a}(x_3-b)^2$, where a and b do not depend on x_3:

$$\begin{aligned}
-\frac{(x_1-x_3)^2}{2} - \frac{(x_2-x_3)^2}{2} - \frac{x_3^2}{2} &= -\frac{1}{2}(x_1^2 - 2x_1x_3 + x_3^2 + x_2^2 - 2x_2x_3 + x_3^2 + x_3^2) \\
&= -\frac{1}{2}(3x_3^2 - 2x_3(x_1+x_2) + x_1^2 + x_2^2) \\
&= -\frac{3}{2}\left(x_3^2 - 2x_3\frac{x_1+x_2}{3} + \frac{x_1^2+x_2^2}{3}\right) \\
&= -\frac{3}{2}\left(x_3 - \frac{x_1+x_2}{3}\right)^2 - \frac{3}{2}\left(\frac{x_1^2+x_2^2}{3} - \frac{(x_1+x_2)^2}{9}\right) \\
&= -\frac{1}{2a}(x_3-b)^2 - \frac{3}{2}\left(\frac{x_1^2+x_2^2}{3} - \frac{(x_1+x_2)^2}{9}\right),
\end{aligned}$$

where $a = 1/3$ and $b = \frac{x_1+x_2}{3}$. This shows that for fixed x_1 and x_2, the joint pdf $f_X(x_1,x_2,x_3)$ as a function of x_3 is a constant, say c, times a $N(b,a)$ density for x_3. Integrating over all x_3 leaves just the constant c, which depends on x_1 and x_2. Then c considered as a function of x_1 and x_2 is the joint marginal pdf of (X_1, X_2), which is

$$\begin{aligned}
f_{X_1,X_2}(x_1,x_2) &= \frac{\sqrt{a}}{2\pi}\exp\left\{-\frac{3}{2}\Big(\frac{x_1^2+x_2^2}{3} - \frac{(x_1+x_2)^2}{9}\Big)\right\} \\
&= \frac{1}{2\pi\sqrt{3}}\exp\left\{-\frac{1}{6}(3x_1^2 + 3x_2^2 - x_1^2 - 2x_1x_2 - x_2^2)\right\} \\
&= \frac{1}{2\pi\sqrt{3}}\exp\left\{-\frac{1}{3}(x_1^2 - x_1x_2 + x_2^2)\right\},
\end{aligned}$$

which is valid for $-\infty < x_1, x_2 < \infty$. For the marginal pdf of X_1 we integrate $f_{X_1,X_2}(x_1,x_2)$ over all x_2. We write the exponent as

$$-\frac{1}{3}(x_1^2 - x_1x_2 + x_2^2) = -\frac{1}{3}\left(x_2^2 - 2\frac{x_1x_2}{2} + x_1^2\right) = -\frac{1}{3}\left(\left(x_2 - \frac{x_1}{2}\right)^2 + x_1^2 - \frac{x_1^2}{4}\right)$$

So, for fixed x_1, $f_{X_1,X_2}(x_1,x_2)$ is a constant times a $N(\frac{x_1}{2}, \frac{3}{2})$ density for x_2. Integrating out x_2 leaves the constant, which as a function of x_1 gives the marginal pdf of X_1:

$$f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}\sqrt{3}}\sqrt{\frac{3}{2}}\exp\left\{-\frac{1}{3}\left(x_1^2 - \frac{x_1^2}{4}\right)\right\} = \frac{1}{\sqrt{2\pi}\sqrt{2}}\exp\left\{-\frac{1}{2(2)}x_1^2\right\},$$

i.e., X_1 has a $N(0,2)$ distribution.

Student Number Name

4. (4 marks) An urn contains 2 balls numbered '1', 2 balls numbered '2', and 10 balls numbered '3'. Seven balls are drawn at random from the urn, without replacement. Let X_i be the number of balls in the sample that are numbered 'i', for $i = 1, 2, 3$. Find $E[X_3]$.

Solution: We first get the marginal pmf of X_3. The possible values of X_3 are 3, 4, 5, 6 and 7. Relabel the balls numbered '1' and '2' as '0', so the urn contains 4 balls labelled '0' and 10 balls labelled '3'. Let X_0 denote the number of balls in the sample that are labelled '0'. For $x_3 \in \{3, 4, 5, 6, 7\}$,

$$P(X_3 = x_3) = P(X_3 = x_3, X_0 = 7 - x_3) = \frac{\binom{10}{x_3}\binom{4}{7-x_3}}{\binom{14}{7}}.$$

Then

$$\begin{aligned}
E[X_3] &= \sum_{x_3=3}^{7} x_3 \frac{\binom{10}{x_3}\binom{4}{7-x_3}}{\binom{14}{7}} \\
&= \frac{1}{\binom{14}{7}} \left[3\binom{10}{3}\binom{4}{4} + 4\binom{10}{4}\binom{4}{3} + 5\binom{10}{5}\binom{4}{2} + 6\binom{10}{6}\binom{4}{1} + 7\binom{10}{7}\binom{4}{0} \right] \\
&= \frac{1}{3432}\Big(3(120)(1) + 4(210)(4) + 5(252)(6) + 6(210)(4) + 7(120)(1)\Big) \\
&= \frac{17160}{3432} = 5.
\end{aligned}$$

Student Number Name

5. (5 marks) We have seen in the class lecture notes that if $X_1, \ldots, X_n$ are independent then $g_1(X_1), \ldots, g_n(X_n)$ are also independent, where $g_1(\cdot), \ldots, g_n(\cdot)$ are arbitrary real-valued functions. Hence, it follows that $E[g_1(X_1) \ldots g_n(X_n)] = E[g_1(X_1)] \ldots E[g_n(X_n)]$ (assuming the expectations exist). Conversely, show that if

$$E[g_1(X_1) \ldots g_n(X_n)] = E[g_1(X_1)] \ldots E[g_n(X_n)]$$

holds for all functions $g_1, \ldots, g_n$ for which the expectations exist, then $X_1, \ldots, X_n$ are mutually independent. *Hint:* Consider functions which are indicators of sets.

Solution: Suppose that $E[g_1(X_1) \ldots g_n(X_n)] = E[g_1(X_1)] \ldots E[g_n(X_n)]$ for all functions $g_1, \ldots, g_n$ for which the expectations exist. Let $A_1, \ldots, A_n$ be arbitrary events on the real line. We wish to show that

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \times \ldots \times P(X_n \in A_n).$$

Let $g_i(X_i) = I_{A_i}(X_i)$, the indicator that X_i is in A_i, for $i = 1, \ldots, n$; that is,

$$I_{A_i}(X_i) = \begin{cases} 1 & \text{if } X_i \in A_i \\ 0 & \text{if } X_i \notin A_i. \end{cases}$$

Then noting that $I_{A_1}(X_1) \ldots I_{A_n}(X_n) = 1$ if and only if the event $\{X_1 \in A_1, \ldots, X_n \in A_n\}$ occurs, we have

$$\begin{aligned} P(X_1 \in A_1, \ldots, X_n \in A_n) &= E[I_{A_1}(X_1) \ldots I_{A_n}(X_n)] \\ &= E[g_1(X_1) \ldots g_n(X_n)] \\ &= E[g_1(X_1)] \ldots E[g_n(X_n)] \\ &= E[I_{A_1}(X_1)] \ldots E[I_{A_n}(X_n)] \\ &= P(X_1 \in A_1) \ldots P(X_n \in A_n). \end{aligned}$$

Since $A_1, \ldots, A_n$ were arbitrary, this implies that $X_1, \ldots, X_n$ are mutually independent.