Homework 3

Independence

Jaiden Atterbury

Instructions

Please answer the following questions in the order in which they are posed. Add a few empty lines below each and write your answers there. **Focus on answering in complete sentences and show work whether we ask for it or not**. You will also need scratch paper/pen to work out the answers before typing it.

For help with formatting documents in RMarkdown, please consult R Markdown: The Definitive Guide. Another option is to search using Google.

Exercises

1. The random variables X and Y are independent, each taking the values 1, 2 or 3. Complete the following table of the joint PMF. Show your work for each entry below the table.

$y \backslash x$	1	2	3
1	0.03	0.04	0.03
2	0.15	(b)	(c)
3	(a)	(d)	(e)

Solving for (a):

$$
\begin{aligned}
f(1,1) &= f_1(1) \times f_2(1) = 0.03 \\
&= f_1(1) \times (0.03 + 0.04 + 0.03) = 0.03 \\
&= f_1(1) \times 0.1 = 0.03 \\
&= f_1(1) = \frac{0.03}{0.1} = 0.3 \\
f_1(1) &= 0.03 + 0.15 + a = 0.3 \\
a &= 0.12
\end{aligned}
$$

Now, in order to find (b)-(e), we must start by finding the marginal values of X and Y.

Solving for $\mathbf{f_1(x)}$:

$$
\begin{aligned}
f(2,1) &= f_1(2) \times f_2(1) \\
&= (0.04 + b + d)(0.1) = 0.04 \\
&= 0.04 + b + d = 0.4 \\
&= b + d = 0.36 \\
\therefore f_1(2) &= 0.04 + 0.36 = 0.4 \\
\therefore f_1(3) &= 1 - (0.3 + 0.4) = 0.3
\end{aligned}
$$

Solving for $\mathbf{f_2(y)}$:

$$\begin{aligned}
f(1,2) &= f_1(1) \times f_2(2) \\
&= (0.3)(0.15 + b + c) = 0.04 \\
&= 0.15 + b + c = 0.5 \\
&= b + c = 0.35 \\
\therefore f_2(2) &= 0.15 + 0.35 = 0.5 \\
\therefore f_2(3) &= 1 - (0.1 + 0.5) = 0.4
\end{aligned}$$

We can now use the fact that when X and Y are independent $f(y|x) = \frac{f(x,y)}{f_1(x)} = f_2(y)$, along with the fact that the the rows and columns must add up to their corresponding marginal values to find (b)-(e).

Solving for (b):

$$\begin{aligned}
f(y|x) &= \frac{f(x,y)}{f_1(x)} = f_2(y) \quad \text{Since } X \text{ and } Y \text{ are independent} \\
f(y=2|x=2) &= \frac{f(2,2)}{f_1(2)} = f_2(2) \\
&= \frac{b}{0.4} = 0.5 \\
b &= 0.2
\end{aligned}$$

Solving for (c):

$$\begin{aligned}
b + c &= 0.35 \\
0.2 + c &= 0.35 \\
c &= 0.15
\end{aligned}$$

Solving for (d):

$$\begin{aligned}
b + d &= 0.36 \\
0.2 + d &= 0.36 \\
d &= 0.16
\end{aligned}$$

Solving for (e):

$$\begin{aligned}
0.4 &= 0.12 + 0.16 + e \\
0.4 &= 0.28 + e \\
e &= 0.12
\end{aligned}$$

Final table:

$y \backslash x$	1	2	3
1	0.03	0.04	0.03
2	0.15	0.2	0.15
3	0.12	0.16	0.12

2. Suppose the probabilistic behavior of $\langle X, Y \rangle$ is governed by the joint PDF

$$f(x,y) = cx^2\, y^4\, e^{-y}\, e^{-\frac{x}{2}} \quad 0 < x,\ 0 < y.$$

a. Determine c. Show work.

Hint: you will need to remember properties of the Gamma function $\Gamma(k)$ which is defined for $k > 0$ as:

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$$

In order for $f(x,y)$ to be a valid pdf, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\ dy\ dx = 1$. This will be a key fact in solving for c.

Solving for c:

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\ dy\ dx \\
&= c \int_0^\infty \int_0^\infty x^2 y^4 e^{-y} e^{\frac{-x}{2}}\ dy\ dx \\
&= c \int_0^\infty x^2 e^{\frac{-x}{2}}\ dx \int_0^\infty y^4 e^{-y}\ dy \\
&= c\Gamma(5) \int_0^\infty x^2 e^{\frac{-x}{2}}\ dx \quad \text{Since } \Gamma(5) = \int_0^\infty x^4 e^{-x} dx \\
&= c \cdot 2\Gamma(5) \int_0^\infty (2u)^2 e^{-u}\ du \quad \text{Let u} = \frac{1}{2}\text{x} \\
&= c \cdot 8\Gamma(5) \int_0^\infty u^2 e^{-u}\ du \\
&= c \cdot 8\Gamma(5)\Gamma(3) \quad \text{Since } \Gamma(3) = \int_0^\infty x^2 e^{-x} dx \\
c &= \frac{1}{8\Gamma(5)\Gamma(3)} \\
c &= \frac{1}{8 \cdot 4! \cdot 2!} \quad \text{since } \Gamma(k) = (k-1)! \\
c &= \frac{1}{384}
\end{aligned}$$

b. Are X and Y are independent? Explain.

In order to prove that X and Y are independent, we must show that $f(x,y) = f_1(x) \times f_2(y)$ for all x,y.

Solving for $\mathbf{f_1(x)}$:

$$\begin{aligned}
f_1(x) &= \int_{-\infty}^{\infty} f(x,y)\ dy \\
&= \int_{0}^{\infty} \frac{1}{384} x^2 y^4 e^{-y} e^{\frac{-x}{2}}\ dy \\
&= \frac{1}{384} x^2 e^{\frac{-x}{2}} \int_{0}^{\infty} y^4 e^{-y}\ dy \\
&= \frac{\Gamma(5)}{384} x^2 e^{\frac{-x}{2}} \quad \text{HW3 Problem 2a.} \\
&= \frac{1}{16} x^2 e^{\frac{-x}{2}},\ 0 < x
\end{aligned}$$

Solving for $\mathbf{f_2(y)}$:

$$\begin{aligned}
f_2(y) &= \int_{-\infty}^{\infty} f(x,y)\ dx \\
&= \int_{0}^{\infty} \frac{1}{384} x^2 y^4 e^{-y} e^{\frac{-x}{2}}\ dx \\
&= \frac{1}{384} y^4 e^{-y} \int_{0}^{\infty} x^2 e^{\frac{-x}{2}}\ dx \\
&= \frac{8\Gamma(3)}{384} y^4 e^{-y} \quad \text{HW3 Problem 2a.} \\
&= \frac{1}{24} y^4 e^{-y},\ 0 < y
\end{aligned}$$

Putting it all together:

$$\begin{aligned}
f_1(x) \times f_2(y) &= (\frac{1}{16} x^2 e^{\frac{-x}{2}})(\frac{1}{24} y^4 e^{-y}) \\
&= \frac{1}{384} x^2 e^{\frac{-x}{2}} y^4 e^{-y},\ x < 0\ y < 0 \\
&= f(x,y)
\end{aligned}$$

Therefore, X and Y are independent.

3. Two friends - let's call them Henry and Vincent - agree to meet at Tully's for coffee. Suppose that the random variables

$$X = \text{the time that Henry arrives at Tully's and}$$

$$Y = \text{the time that Vincent arrives at Tully's}$$

are independent uniform random variables on the interval $[5, 6]$. (The units are hours after noon)

a. Calculate the probability that both of them arrive before 5:30 PM.

Since $X \sim Unif(5,6)$ and $Y \sim Unif(5,6)$, it follows from the definition of a Uniform PDF that the PDFs of X and Y are $f_1(x) = \frac{1}{b-a} = \frac{1}{6-5} = 1,\ 5 \le x \le 6$ and $f_2(y) = \frac{1}{b-a} = \frac{1}{6-5} = 1,\ 5 \le y \le 6$. Since we know that X and Y are independent, and we also know the individual/marginal PDFs of X and Y, we can use the fact that if X and Y are independent $f(x,y) = f_1(x) \times f_2(y)$ to find the joint PDF of X and Y. Thus the joint PDF of X and Y is $f(x,y) = 1,\ 5 \le x \le 6$ and $5 \le y \le 6$.

Furthermore, since 5:30PM is the same as 5.5 in terms of the given interval for X and Y, the probability that both of them arrive before 5:30 PM translates to $P(X < 5.5, Y < 5.5)$.

Solving for $\mathbf{P(X < 5.5, Y < 5.5)}$:

$$\begin{aligned}
P(X < 5.5, Y < 5.5) &= \int_5^{5.5}\int_5^{5.5} f(x,y)\ dy\ dx \\
&= \int_5^{5.5}\int_5^{5.5} dy\ dx \\
&= \int_5^{5.5} dx \int_5^{5.5} dy \\
&= [x]_5^{5.5} \cdot [y]_5^{5.5} \\
&= \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{4}
\end{aligned}$$

b. If neither of them is willing to wait more than 10 minutes for the other to show up, what is the probability they have a coffee together? That is, what is the probability they arrive within 10 minutes of each other.

Hint: Here you want to find the probability that $|X - Y|$ is less than some number.

If neither of them is willing to wait more than 10 minutes for the other to show up, and 10 minutes is equivalent to $\frac{1}{6}$ in terms of the given interval, then the probability they have a coffee together is the same as saying the probability they arrive within 10 minutes of each other. In terms of the random variables, this translates to $P(|X - Y| \le \frac{1}{6}$.

In terms of x, the expression $|x - y| \le \frac{1}{6}$ translates to $y \ge x - \frac{1}{6}$ and $y \le x + \frac{1}{6}$. When taking into account that $5 \le x \le 6$ and $5 \le y \le 6$, it turns out that there are 3 distinct regions that we must integrate over in order to find this probability. We will call them $R_1, R_2,$ and R_3. In particular, $P(|X - Y| \le \frac{1}{6}) = R_1 + R_2 + R_3$.

Calculating $\mathbf{R_1}$:

$$\begin{aligned}
R_1 &= \int_5^{\frac{31}{6}}\int_5^{x+\frac{1}{6}} dy\ dx \\
&= \int_5^{\frac{31}{6}} x + \frac{1}{6} - 5\ dx \\
&= \int_5^{\frac{31}{6}} x - \frac{29}{6}\ dx \\
&= \left[\frac{1}{2}x^2 - \frac{29}{6}x\right]_5^{\frac{31}{6}} \\
&= \frac{1}{24}
\end{aligned}$$

Calculating $\mathbf{R_2}$:

$$
\begin{aligned}
R_2 &= \int_{\frac{31}{6}}^{\frac{35}{6}} \int_{x-\frac{1}{6}}^{x+\frac{1}{6}} dy\ dx \\
&= \int_{\frac{31}{6}}^{\frac{35}{6}} (x - \frac{1}{6}) - (x + \frac{1}{6})\ dx \\
&= \int_{\frac{31}{6}}^{\frac{35}{6}} \frac{1}{3}\ dx \\
&= \left[\frac{1}{3}x\right]_{\frac{31}{6}}^{\frac{35}{6}} \\
&= \frac{2}{9}
\end{aligned}
$$

Calculating $\mathbf{R_3}$:

$$
\begin{aligned}
R_3 &= \int_{\frac{35}{6}}^{6} \int_{x-\frac{1}{6}}^{6} dy\ dx \\
&= \int_{\frac{35}{6}}^{6} 6 - (x - \frac{1}{6})\ dx \\
&= \int_{\frac{35}{6}}^{6} \frac{35}{6} - x\ dx \\
&= \left[\frac{35}{6}x - \frac{1}{2}x^2\right]_{\frac{35}{6}}^{6} \\
&= \frac{1}{24}
\end{aligned}
$$

Putting it all together:

$$
\begin{aligned}
P(|X - Y| \leq \frac{1}{6}) &= R_1 + R_2 + R_3 \\
&= \frac{1}{24} + \frac{2}{9} + \frac{1}{24} \\
&= \frac{11}{36}
\end{aligned}
$$

Thus, the probability that Henry and Vincent end up meeting each other at Tully's for coffee is $\frac{11}{36}$

4. An individual makes 100 check transactions between receiving his December and his January bank statements. Rather than subtracting the amounts exactly, he rounds off each checkbook entry to the nearest dollar. Let X_i denote the round off error on the *ith* check. A reasonable assumption is that

$$X_i \sim Unif\left(-\frac{1}{2}, \frac{1}{2}\right)$$

independently of each other.

Use the Bienaymé-Chebyshev inequality to get an upper bound for the probability that the accumulated error after his 100 transactions is \$5 or more. Show work.

Problem setup:
To start off this problem, let Z denote the accumulated error of this individuals 100 transactions. Thus, the probability that the accumulated error after his 100 transactions is 5 or more, is equivalent to the probability statement: $P(|Z| \geq 5)$. However, we aren't just looking for this probability, instead we are using the Bienaymé-Chebyshev inequality to get an upper bound for this probability. The Bienaymé-Chebyshev inequality states $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$. With that said, in order to find this upper bound for Z we need to find E[Z] and SD[Z] in order to solve for k.

Since Z denotes the accumulated error of this individuals 100 transactions and each X_i denotes the round off error on the *ith* check, then it follows that $Z = X_1 + X_2 + \cdots + X_{100}$. Thus, we can use the fact that $X_i \sim Unif(-\frac{1}{2}, \frac{1}{2})$ and that all of the X_i's are independent to find E[Z] and SD[Z].

Finding $\mathbf{E[X_i]}$, $\mathbf{Var[X_i]}$:
Since $X_i \sim Unif(-\frac{1}{2}, \frac{1}{2})$, it follows from Example 12.1 that $E[X_i] = \frac{\frac{-1}{2}+\frac{1}{2}}{2} = \frac{0}{2} = 0$ and $Var[X_i] = \frac{(\frac{1}{2}-\frac{-1}{2})^2}{12} = \frac{1^2}{12} = \frac{1}{12}$.

Finding E[Z], Var[Z], and SD[Z]:

$$\begin{aligned}
E[Z] &= E[X_1 + X_2 + \cdots + X_{100}] \\
&= E[X_1] + E[X_2] + \cdots + E[X_{100}] \\
&= 0 + 0 + \cdots + 0 \\
&= 100 \cdot 0 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
Var[Z] &= Var[X_1 + X_2 + \cdots + X_{100}] \\
&= Var[X_1] + Var[X_2] + \cdots + Var[X_{100}] \quad \text{Since the } X_i\text{'s are independent} \\
&= \frac{1}{12} + \frac{1}{12} + \cdots + \frac{1}{12} \\
&= 100 \cdot \frac{1}{12} \\
&= \frac{25}{3}
\end{aligned}$$

$$\begin{aligned}
SD[Z] &= \sqrt{Var[Z]} \\
&= \sqrt{Var[X_1 + X_2 + \cdots + X_{100}]} \\
&= \sqrt{\frac{1}{12} + \frac{1}{12} + \cdots + \frac{1}{12}} \\
&= \sqrt{\frac{25}{3}} \\
&= \frac{5\sqrt{3}}{3}
\end{aligned}$$

Finding the upper bound:
Since the Bienaymé-Chebyshev inequality states $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$, in our case this translates to $P(|Z| \geq k\frac{5\sqrt{3}}{3}) \leq \frac{1}{k^2}$. Since we know that $k\frac{5\sqrt{3}}{3} = 5$ and $-k\frac{5\sqrt{3}}{3} = -5$ all we have to do is solve for k then we will have our answer. Solving for k simply shows us that $k = 5 \cdot \frac{3}{5\sqrt{3}} = \frac{3}{\sqrt{3}} = \sqrt{3}$.

Thus, it follows that an upper bound for the probability that the accumulated error after his 100 transactions is 5 or more is $\frac{1}{\sqrt{3}^2} = \frac{1}{3}$. This translates to the probability statement $P(|Z| \geq 5) \leq \frac{1}{3}$.