Statistics for Data Science - 2

Week 3 Practice Assignment Solution

Multiple random variables

1. Let X and Y be two random variables with joint distribution given in Table 3.1.P, where a and b are two unknown values.

$Y \backslash X$	0	1	2
0	$\frac{1}{12}$	$\frac{3}{12}$	a
1	$\frac{2}{12}$	b	$\frac{1}{12}$
2	$\frac{3}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Table 3.1.P: Joint distribution of X and Y.

i) Find $P(Y = 1)$.

a) $\frac{4}{12}$

b) $\frac{3}{12}$

c) $\frac{5}{12}$

d) $\frac{1}{12}$

Solution:
We know that, $\sum\limits_{x \in T_X,\ y \in T_Y} f_{XY}(x, y) = 1$

$\Rightarrow \frac{1}{12} + \frac{3}{12} + a + \frac{2}{12} + b + \frac{1}{12} + \frac{3}{12} + \frac{1}{12} + \frac{1}{12} = 1$

$\Rightarrow a + b = 0$

Since a and b cannot take negative values $\Rightarrow a = b = 0$.

Now,

$$
\begin{aligned}
P(Y=1) &= \sum_{x \in T_X} f_{XY}(x,1) \\
&= \frac{2}{12} + b + \frac{1}{12} \\
&= \frac{3}{12} + 0 \\
&= \frac{3}{12}
\end{aligned}
$$

ii) Find $P(Y=1 \mid X=2)$.

a) $\frac{1}{12}$

b) $\frac{1}{4}$

c) $\frac{1}{3}$

d) $\frac{1}{2}$

Solution:

$$
\begin{aligned}
P(Y=1 \mid X=2) &= \frac{P(Y=1, X=2)}{P(X=2)} \\
&= \frac{\frac{1}{12}}{a + \frac{1}{12} + \frac{1}{12}} \\
&= \frac{1}{2}
\end{aligned}
$$

iii) Find $P(X=0, Y \geq 1)$.

a) $\frac{4}{12}$

b) $\frac{3}{12}$

c) $\frac{5}{12}$

d) $\frac{1}{12}$

Solution:

$$
\begin{aligned}
P(X=0, Y \geq 1) &= P(X=0, Y=1) + P(X=0, Y=2) \\
&= \frac{2}{12} + \frac{3}{12} \\
&= \frac{5}{12}
\end{aligned}
$$

2. Let X and Y be two independent discrete random variables with CDFs F_X and F_Y, respectively. Define another random variable $Z = \min(X, Y)$, then the CDF of Z is

   a) $\min(F_X, F_Y)$
   b) $F_X F_Y$
   c) $F_X + F_Y + F_X F_Y$
   d) $F_X + F_Y - F_X F_Y$

   Solution:

$$
\begin{aligned}
F_Z(z) = P(Z \leq z) &= P(\min(X, Y) \leq z) \\
&= 1 - P(\min(X, Y) > z) \\
&= 1 - P(X > z, Y > z)
\end{aligned}
$$

   Since X and Y are two independent discrete random variables,
   $P(X > z, Y > z) = P(X > z)P(Y > z)$

$$
\begin{aligned}
\Rightarrow F_Z(z) &= 1 - P(X > z)P(Y > z) \\
&= 1 - [(1 - P(X \leq z))(1 - P(Y \leq z))] \\
&= 1 - [(1 - F_X(z))(1 - F_Y(z))] \\
&= F_X(z) + F_Y(z) - F_X(z)F_Y(z)
\end{aligned}
$$

3. Let X and Y be two independent random variables with PMFs

$$
f_X(k) = f_Y(k) = \begin{cases} \frac{1}{6} & \text{for } k = 1, 2, 3, 4, 5, 6. \\ 0 & \text{otherwise} \end{cases}
$$

   Define $Z = X - Y$. Find the value of $f_Z(3)$.

   a) $\frac{4}{12}$
   b) $\frac{3}{12}$

c) $\frac{5}{12}$

d) $\boxed{\frac{1}{12}}$

Solution:

$$\begin{aligned} f_Z(3) = P(Z=3) &= P(X-Y=3) \\ &= P(X=4, Y=1) + P(X=5, Y=2) + P(X=6, Y=3) \end{aligned}$$

Given that X and Y are two independent random variables.
$\Rightarrow P(X=x, Y=y) = P(X=x)P(Y=y)$ for all (x, y).

$$\begin{aligned} f_Z(3) &= P(X=4, Y=1) + P(X=5, Y=2) + P(X=6, Y=3) \\ &= P(X=4)P(Y=1) + P(X=5)P(Y=2) + P(X=6)P(Y=3) \\ &= \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} \\ &= \frac{3}{36} \\ &= \frac{1}{12} \end{aligned}$$

4. Let $X \sim$ Geometric(p) and $Y \sim$ Geometric(p) be independent and let $Z = X + Y$. Determine the values of p for which $P(Z = 26) > P(Z = 25)$.

   a) $p > 0.02$
   b) $\boxed{p < 0.04}$
   c) $p > 0.15$
   d) $p < 0.30$
   e) $p = 0.05$

Solution:
If $X \sim$ Geometric(p) and $Y \sim$ Geometric(p) are two independent random variables and $Z = X + Y$, then
$P(Z = n) = (n-1)p^2(1-p)^{n-2}$ (try derivation by yourself)
We have to find the value of p for which $P(Z = 26) > P(Z = 25)$.
$P(Z = 26) = (26-1)p^2(1-p)^{26-2}$ and $P(Z = 25) = (25-1)p^2(1-p)^{25-2}$
Comparing both, we will get
$25p^2(1-p)^{24} > 24p^2(1-p)^{23}$
$\Rightarrow 25(1-p) > 24$
$\Rightarrow 1 - p > \frac{24}{25}$
$\Rightarrow p < 0.04$

5. Let $X \sim \text{Uniform}(\{1, 2, 3, 4, 5, 6\})$ and let Y be the number of times 2 occurs in X throws of a fair die. Choose the **incorrect** option(s) among the following.

   a) $P(Y = 2 \mid X = 2) = \dfrac{1}{6}$

   b) $P(Y = 2 \mid X = 4) = \dfrac{5^2}{6^3}$

   c) $P(Y = 5 \mid X = 6) = \dfrac{5}{6^5}$

   d) $P(Y = 6 \mid X = 5) = \dfrac{5}{6^6}$

   Solution:

$$\begin{aligned} P(Y = 2 \mid X = 2) &\sim \text{Bin}(2, 1/6), \quad Y \text{takes values in} \{0, 1, 2\} \\ &= {}^2C_2(\frac{1}{6})^2(\frac{5}{6})^0 \\ &= \frac{1}{36} \end{aligned}$$

$$\begin{aligned} P(Y = 2 \mid X = 4) &\sim \text{Bin}(4, 1/6) \\ &= {}^4C_2(\frac{1}{6})^2(\frac{5}{6})^2 \\ &= \frac{5^2}{6^3} \end{aligned}$$

$$\begin{aligned} P(Y = 5 \mid X = 6) &\sim \text{Bin}(6, 1/6) \\ &= {}^6C_5(\frac{1}{6})^5(\frac{5}{6})^1 \\ &= \frac{5}{6^5} \end{aligned}$$

$$\begin{aligned} P(Y = 6 \mid X = 5) &\sim \text{Bin}(5, 1/6) \\ &= 0 \end{aligned}$$

6. Let the random variables X and Y each have range $\{1,\ 2,\ 3\}$. The following formula gives the joint PMF

$$P(X = i, Y = j) = \frac{i + 2j}{c},$$

   where c is an unknown value. Find $P(1 \leq X \leq 3, 1 < Y \leq 3)$.

a) $\frac{5}{9}$

b) $\frac{7}{9}$

c) $\frac{2}{9}$

d) $\frac{4}{9}$

Solution:
We know that, $\sum\limits_{x \in T_X,\ y \in T_Y} P(X = x, Y = y) = 1$

$\Rightarrow P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) + P(X = 3, Y = 1) + P(X = 3, Y = 2) + P(X = 3, Y = 3) = 1$

$\Rightarrow \frac{3}{c} + \frac{5}{c} + \frac{7}{c} + \frac{4}{c} + \frac{6}{c} + \frac{8}{c} + \frac{5}{c} + \frac{7}{c} + \frac{9}{c} = 1$

$\Rightarrow c = 54$

Now,

$$\begin{aligned} P(1 \leq X \leq 3, 1 < Y \leq 3) = {} & P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 2, Y = 2) \\ & + P(X = 2, Y = 3) + P(X = 3, Y = 2) + P(X = 3, Y = 3) \end{aligned}$$

$$\begin{aligned} \Rightarrow P(1 \leq X \leq 3, 1 < Y \leq 3) &= \frac{1}{c}[5 + 7 + 6 + 8 + 7 + 9] \\ &= \frac{42}{54} \\ &= \frac{7}{9} \end{aligned}$$

7. The joint PMF of the random variables X and Y is given in Table 3.2.P.

Y \ X	1	2	3
1	k	k	$2k$
2	$2k$	0	$4k$
3	$3k$	k	$6k$

Table 3.2.P: Joint distribution of X and Y.

Consider the random variable $Z = X^2 Y$.
i) Find the range of $Z \mid Y = 2$.

a) $\{1, 4, 9\}$

b) $\{4, 8, 18\}$

c) $\{1, 9\}$

d) $\{2, 18\}$

e) $\{2, 8, 18\}$

Solution:
We know that, $\sum\limits_{x\in T_X,\ y\in T_Y} P(X = x, Y = y) = 1$
$\Rightarrow k + k + 2k + 2k + 0 + 4k + 3k + k + 6k = 1$
$\Rightarrow k = \frac{1}{20}$
When $Y = 2, P(X = 2, Y = 2) = 0$. So for the range we will not consider the pair (2, 2).
Since $Z = X^2Y$, the range of $Z \mid Y = 2$ will be $\{1^2 \times 2, 3^2 \times 2\}$ which is equal to $\{2, 18\}$.

ii) Find the value of $P(Z = 18 \mid Y = 2)$.

a) $\frac{1}{3}$

b) $\frac{2}{3}$

c) $\frac{3}{4}$

d) $\frac{1}{4}$

Solution:

$$\begin{aligned}
P(Z = 18 \mid Y = 2) &= \frac{P(Z = 18, Y = 2)}{P(Y = 2)} \\
&= \frac{P(X = 3, Y = 2)}{P(X = 1, Y = 2) + P(X = 3, Y = 2)} \\
&= \frac{4k}{2k + 4k} \\
&= \frac{2}{3}
\end{aligned}$$

8. The following options gives the joint PMF of the random variables X and Y. If the random variables X and Y are independent, then which of the following option(s) can be the joint PMF of X and Y?

Y / X	0	1	2
0	0.01	0	0
1	0.09	0.09	0
2	0	0	0.81

a)

Y / X	0	1	2
0	0.06	0.18	0.12
1	0.04	0.12	0.48

b)

Y / X	0	1	2
0	$\frac{1}{12}$	$\frac{1}{24}$	$\frac{1}{24}$
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{8}$
2	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{12}$

c)

Y / X	0	1	2
0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$
1	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{3}{10}$

d)

$X \backslash Y$	0	1
0	0.10	0.15
1	0.20	0.30
2	0.10	0.15

e)

Solution:
In option a)
$P(X=0, Y=1) = 0$ but $P(X=0) = 0.01+0+0 = 0.01$ and $P(Y=1) = 0+0.09+0 = 0.09$
$\Rightarrow P(X=0, Y=1) \neq P(X=0)P(Y=1)$
Therefore, option (a) cannot be the joint PMF of X and Y.

In option b)
$P(X=0, Y=0) = 0.06$ but $P(X=0) = 0.06+0.18+0.12 = 0.36$ and $P(Y=0) = 0.06+0.04 = 0.10$
$\Rightarrow P(X=0, Y=0) = 0.06 \neq 0.036 = P(X=0)P(Y=0)$
Therefore, option (b) cannot be the joint PMF of X and Y.

In option c)
$P(X=1, Y=0) = 1/6$ but $P(X=1) = 1/6+1/12+1/8 = 3/8$ and $P(Y=0) = 1/12+1/6+1/4 = 1/2$
$\Rightarrow P(X=1, Y=0) = 1/6 \neq 3/16 = P(X=1)P(Y=0)$
Therefore, option (c) cannot be the joint PMF of X and Y.

In option d)
$P(X=0, Y=1) = 1/5$ but $P(X=0) = 1/10+1/5+1/5 = 1/2$ and $P(Y=1) = 1/5+1/10 = 3/10$
$\Rightarrow P(X=0, Y=1) = 1/5 \neq 3/20 = P(X=0)P(Y=1)$
Therefore, option (d) cannot be the joint PMF of X and Y.

In option e)
For every $(x, y), P(X=x, Y=y) = P(X=x)P(Y=y)$ (check yourself)
Hence option (e) is the joint PMF of X and Y.
Answer: e

9. From a sack of fruits containing 3 mangoes, 2 kiwis, and 3 guavas, a random sample of 4 pieces of fruit is selected. If X is the number of mangoes and Y is the number of kiwis in the sample, then find the joint probability distribution of X and Y.

a)

$Y \backslash X$	0	1	2	3
0	0	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{2}{70}$	$\frac{18}{70}$
2	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	0

b)

$Y \backslash X$	0	1	2	3
0	0	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{18}{70}$	$\frac{2}{70}$
2	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	0

c)

$Y \backslash X$	0	1	2	3
0	0	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{18}{70}$	$\frac{2}{70}$
2	$\frac{9}{70}$	$\frac{3}{70}$	$\frac{3}{70}$	0

Y \ X	0	1	2	3
0	0	$\frac{3}{70}$	$\frac{3}{70}$	$\frac{9}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{18}{70}$	$\frac{2}{70}$
2	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	0

d)

Solution:
X is the number of mangoes and Y is the number of kiwis in the sample. The number of mangoes and kiwis in the sack is 3 and 2,respectively.
So X will take values in $\{0, 1, 2, 3\}$ and Y will take values in $\{0, 1, 2\}$ when the random sample of 4 pieces is selected.
$P(X = 0, Y = 0) = P(\text{no mango and no kiwi}) = 0$ (not possible since the number of guava is 3)

$P(X = 0, Y = 1) = P(\text{no mango and one kiwi}) = \dfrac{{}^2C_1 {}^3C_3}{{}^8C_4} = \dfrac{2}{70}$

$P(X = 0, Y = 2) = P(\text{no mango and two kiwis}) = \dfrac{{}^2C_2 {}^3C_2}{{}^8C_4} = \dfrac{3}{70}$

$P(X = 1, Y = 0) = P(\text{one mango and no kiwi}) = \dfrac{{}^3C_1 {}^3C_3}{{}^8C_4} = \dfrac{3}{70}$

$P(X = 1, Y = 1) = P(\text{one mango and one kiwi}) = \dfrac{{}^3C_1 {}^2C_1 {}^3C_2}{{}^8C_4} = \dfrac{18}{70}$

$P(X = 1, Y = 2) = P(\text{one mango and two kiwis}) = \dfrac{{}^3C_1 {}^2C_2 {}^3C_1}{{}^8C_4} = \dfrac{9}{70}$

$P(X = 2, Y = 0) = P(\text{two mangoes and no kiwi}) = \dfrac{{}^3C_2 {}^3C_2}{{}^8C_4} = \dfrac{9}{70}$

$P(X = 2, Y = 1) = P(\text{two mangoes and one kiwi}) = \dfrac{{}^3C_2 {}^2C_1 {}^3C_1}{{}^8C_4} = \dfrac{18}{70}$

$P(X = 2, Y = 2) = P(\text{two mangoes and two kiwis}) = \dfrac{{}^3C_2 {}^2C_2}{{}^8C_4} = \dfrac{3}{70}$

Similarly you can check for other values also.
Answer: b

10. Suppose you flip a fair coin. If the coin lands heads, you roll a fair six-sided die 50 times. If the coin lands tails, you roll the die 51 times. Let X be 1 if the coin lands heads and

0 if the coin lands tails. Let Y be the total number of times you get the number 5 while throwing the dice. Find $P(X = 1|Y = 10)$.

a) $\frac{85}{157}$

b) $\frac{82}{167}$

c) $\frac{72}{157}$

d) $\frac{85}{167}$

Solution:

$$\begin{aligned}
\frac{P(X=1|Y=10)}{P(X=0|Y=10)} &= \frac{P(Y=10|X=1).P(X=1)}{P(Y=10|X=0).P(X=0)} \\
&= \frac{P(Y=10|X=1)}{P(Y=10|X=0)} \qquad \text{[Since } P(X=1)=P(X=0)\text{]} \\
&= \frac{{}^{50}C_{10}(\frac{1}{6})^{10}(\frac{5}{6})^{40}}{{}^{51}C_{10}(\frac{1}{6})^{10}(\frac{5}{6})^{41}} \\
&= \frac{{}^{50}C_{10}}{{}^{51}C_{10}} \times \frac{6}{5} \\
&= \frac{41}{51} \times \frac{6}{5} \\
&= \frac{246}{255}
\end{aligned}$$

$\Rightarrow P(X = 1|Y = 10) = \frac{246}{255} \times P(X = 0|Y = 10)$

Also $P(X = 1|Y = 10) + P(X = 0|Y = 10) = 1$

$\Rightarrow P(X = 1|Y = 10) + \frac{255}{246}P(X = 1|Y = 10) = 1$

$\Rightarrow P(X = 1|Y = 10) = \frac{246}{501} = \frac{82}{167}$

11. Three balls are selected at random from a box containing five red, four blue, three yellow and six green coloured balls. If X, Y and Z are the number of red balls, blue balls and green balls respectively, choose the correct option(s) among the following.

a) $\boxed{P(X=1, Y=0, Z=2) = \frac{25}{272}}$

b) $\boxed{P(X=1, Y=1, Z=1) = \frac{5}{34}}$

c) $P(X=1, Y=0 \mid Z=2) = \frac{1}{4}$

d) $\boxed{P(X=0, Y=0, Z=3) = \frac{5}{204}}$

Solution:

$P(X=1, Y=0, Z=2) = P(\text{one red ball and 2 green balls}) = \frac{{}^5C_1\,{}^6C_2}{{}^{18}C_3} = \frac{25}{272}$

$$
\begin{aligned}
P(X=1, Y=1, Z=1) = P(\text{one red ball, one blue ball and 1 green ball}) &= \frac{{}^5C_1\,{}^4C_1\,{}^6C_1}{{}^{18}C_3} \\
&= \frac{5}{34}
\end{aligned}
$$

$P(X=0, Y=0, Z=3) = P(\text{3 green balls}) = \frac{{}^6C_3}{{}^{18}C_3} = \frac{5}{204}$

And

$$
\begin{aligned}
P(X=1, Y=0 \mid Z=2) = P(\text{one red ball given that two balls are green}) &= \frac{{}^5C_1}{{}^{16}C_1} \\
&= \frac{5}{16}
\end{aligned}
$$