Statistics for Data Science - 2

Week 6 graded Assignment

Solution

1. A person randomly chooses a battery from a store which has 40 batteries of type A and 60 batteries of type B. Battery life of type A and type B batteries are exponentially distributed with average life of 4 years and 6 years, respectively. If the chosen battery lasts for 5 years, what is the probability that the battery is of type A?

   (a) $\dfrac{1}{1+e^{\frac{5}{12}}}$

   (b) $\dfrac{1}{1+e^{\frac{-5}{12}}}$

   (c) $\dfrac{e^{\frac{-4}{5}}}{1+e^{\frac{-6}{5}}}$

   (d) $\dfrac{e^{\frac{-6}{5}}}{1+e^{\frac{-4}{5}}}$

   Solution:
   Define a event X as follows:

$$X = \begin{cases} 1 & \text{If the chosen battery is of type A} \\ 0 & \text{If the chosen battery is of type B} \end{cases}$$

Let Y denote the battery life of the chosen battery.
By the given information, we have
$Y|X = 1 \sim \text{Exp}(\frac{1}{4})$ and

$Y|X = 0 \sim \text{Exp}(\frac{1}{6})$

It implies that

$f_{Y|X=1}(y) = \frac{1}{4}e^{\frac{-y}{4}}; y > 0$ and

$f_{Y|X=0}(y) = \frac{1}{6}e^{\frac{-y}{6}}; y > 0$

Also given that

$P(X = 1) = \dfrac{40}{100} = \dfrac{2}{5}$ and

$$P(X=0) = \frac{60}{100} = \frac{3}{5}$$

To find: $f_{X|Y=5}(1)$. Now,

$$f_{X|Y=5}(1) = \frac{f_{Y|X=1}(5).P(X=1)}{f_Y(5)}$$

$$= \frac{f_{Y|X=1}(5).P(X=1)}{f_{Y|X=1}(5).P(X=1) + f_{Y|X=0}(5).P(X=0)}$$

$$= \frac{\frac{1}{4}e^{\frac{-5}{4}}.\frac{2}{5}}{\frac{1}{4}e^{\frac{-5}{4}}.\frac{2}{5} + \frac{1}{6}e^{\frac{-5}{6}}.\frac{3}{5}}$$

$$= \frac{\frac{1}{10}e^{\frac{-5}{4}}}{\frac{1}{10}e^{\frac{-5}{4}} + \frac{1}{10}e^{\frac{-5}{6}}}$$

$$= \frac{e^{\frac{-5}{4}}}{e^{\frac{-5}{4}} + e^{\frac{-5}{6}}}$$

$$= \frac{1}{1 + e^{\frac{5}{12}}}$$

---

2. Let $Y = XZ + X$, where $X \sim \text{Uniform}\{1,2,3\}$ and $Z \sim \text{Normal}(1,4)$ are independent. Find the value of $f_{X|Y=2}(2)$.

   (a) $\dfrac{3\exp(\frac{1}{8})}{3\exp(\frac{1}{8}) + 6 + 2\exp(\frac{2}{9})}$

   (b) $\dfrac{3\exp(\frac{-1}{8})}{3\exp(\frac{-1}{8}) + 6 + 2\exp(\frac{-2}{9})}$

   (c) $\dfrac{2\exp(\frac{-2}{9})}{3\exp(\frac{-1}{8}) + 6 + 2\exp(\frac{-2}{9})}$

   (d) $\dfrac{6}{3\exp(\frac{-1}{32}) + 6 + 2\exp(\frac{-1}{18})}$

Solution:
Given that $X \sim \text{Uniform}\{1,2,3\}$ and $Z \sim \text{Normal}(1,4)$ are independent.
$Y = XZ + X$
It implies that

$Y|X = 1 = Z + 1 \sim \text{Normal}(2, 4)$
$Y|X = 2 = 2Z + 2 \sim \text{Normal}(4, 16)$
$Y|X = 3 = 3Z + 3 \sim \text{Normal}(6, 36)$

Therefore,
$f_{Y|X=1}(y) = \frac{1}{2\sqrt{2\pi}} \exp\left(\frac{-(y-2)^2}{8}\right)$
$f_{Y|X=2}(y) = \frac{1}{4\sqrt{2\pi}} \exp\left(\frac{-(y-4)^2}{32}\right)$
$f_{Y|X=3}(y) = \frac{1}{6\sqrt{2\pi}} \exp\left(\frac{-(y-6)^2}{72}\right)$

To find: $f_{X|Y=2}(2)$.

$$f_{X|Y=2}(2) = \frac{f_{Y|X=2}(2).f_X(2)}{f_{Y|X=2}(2).f_X(2) + f_{Y|X=1}(2).f_X(1) + f_{Y|X=3}(2).f_X(3)}$$

$$= \frac{\frac{1}{4\sqrt{2\pi}} \exp\left(\frac{-(2-4)^2}{32}\right).\frac{1}{3}}{\frac{1}{4\sqrt{2\pi}} \exp\left(\frac{-(2-4)^2}{32}\right).\frac{1}{3} + \frac{1}{2\sqrt{2\pi}} \exp\left(\frac{-(2-2)^2}{8}\right).\frac{1}{3} + \frac{1}{6\sqrt{2\pi}} \exp\left(\frac{-(2-6)^2}{72}\right).\frac{1}{3}}$$

$$= \frac{\frac{1}{4} \exp\left(\frac{-1}{8}\right)}{\frac{1}{4} \exp\left(\frac{-1}{8}\right) + \frac{1}{2} \exp(0) + \frac{1}{6} \exp\left(\frac{-2}{9}\right)}$$

$$= \frac{3 \exp(\frac{-1}{8})}{3 \exp(\frac{-1}{8}) + 6 + 2 \exp(\frac{-2}{9})}$$

---

3. The joint pdf of two continuous ranodm variables X and Y is given by

$$f_{XY}(x, y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

1. Yes
2. No

Solution:
First we will calculate the marginal densities of X and Y.

For $0 \leq x \leq 1$

$$\begin{aligned}
f_X(x) &= \int_0^1 f_{XY}(x,y)dy \\
&= \int_0^1 4xydy \\
&= 2xy^2\Big|_0^1 \\
&= 2x
\end{aligned}$$

For $0 \leq y \leq 1$

$$\begin{aligned}
f_Y(y) &= \int_0^1 f_{XY}(x,y)dx \\
&= \int_0^1 4xydx \\
&= 2x^2y\Big|_0^1 \\
&= 2y
\end{aligned}$$

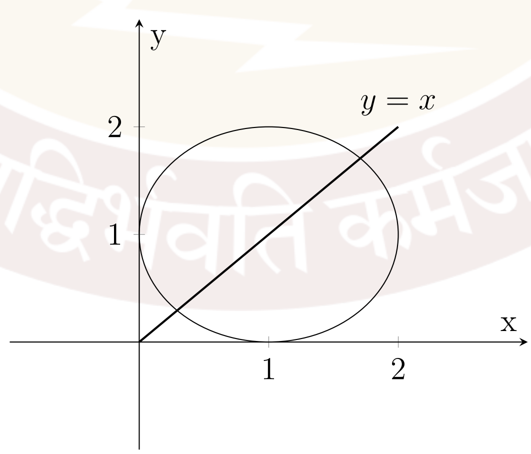
Therefore,

$$f_X(x).f_Y(y) = 4xy = f_{XY}(x,y)$$

It implies that X and Y are independent random variables.

---

4. Let $(X,Y) \sim \text{Uniform}(D)$, where $D = \{(x,y) : (x-1)^2 + (y-1)^2 \leq 1\}$. Calculate $P(X \geq Y)$.
Solution:



The region $X \geq Y$ will be the lower half part of the circle.

Therefore,

$$\begin{aligned} P(X \geq Y) &= \frac{\text{Area of lower half circle}}{\text{Area of the circle}} \\ &= \frac{\pi(1)^2/2}{\pi(1)^2} \\ &= \frac{1}{2} \end{aligned}$$

---

5. Let $(X, Y) \sim \text{Uniform}(D)$, where $D = \{(x, y) : y \leq 2x, 0 < x < 1, 0 < y < 2\} \cup [1, 2] \times [0, 2]$. Find the marginal density of X.

(a)

$$f_X(x) = \begin{cases} \dfrac{2x}{3} + \dfrac{2}{3} & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(b)

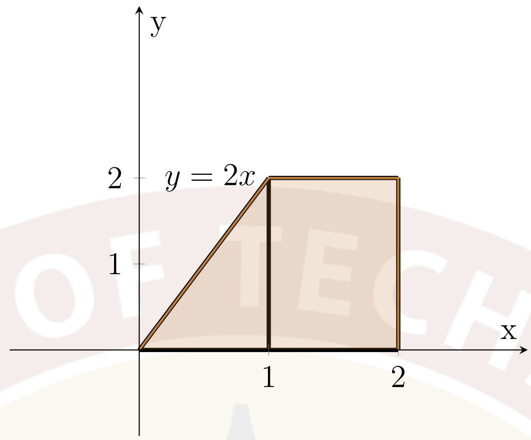
$$f_X(x) = \begin{cases} \dfrac{2x}{3} + \dfrac{1}{3} & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(c)

$$f_X(x) = \begin{cases} \frac{2x}{3} & 0 \leq x \leq 1 \\ \frac{2}{3} & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

(d)

$$f_X(x) = \begin{cases} \frac{2x}{3} & 0 \leq x \leq 1 \\ \frac{1}{3} & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$



D denotes the area of the support(X, Y).
Area of $D = \frac{1}{2} \times 1 \times 2 + 1 \times 2 = 3$

Since $(X, Y) \sim \text{Uniform}(D)$, it implies that

$$f_{XY}(x, y) = \frac{1}{3}, \qquad x, y \in D$$

We know that $f_X(x) = \int f_{XY}(x, y)dy$

For $0 < x < 1$

$$\begin{aligned} f_X(x) &= \int_0^{2x} \frac{1}{3} dy \\ &= \frac{1}{3} y \Big|_0^{2x} \\ &= \frac{2x}{3} \end{aligned}$$

For $1 < x < 2$

$$\begin{aligned} f_X(x) &= \int_0^{2} \frac{1}{3} dy \\ &= \frac{1}{3} y \Big|_0^{2} \\ &= \frac{2}{3} \end{aligned}$$

Therefore, marginal density of X is given by

$$f_X(x) = \begin{cases} \frac{2x}{3} & 0 \le x \le 1 \\ \frac{2}{3} & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

6. The joint pdf of two random variables X and Y is given by

$$f_{XY}(x,y) = \begin{cases} 24xy & 0 \le x \le 1, 0 \le y \le 1, x+y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Choose the correct option(s).

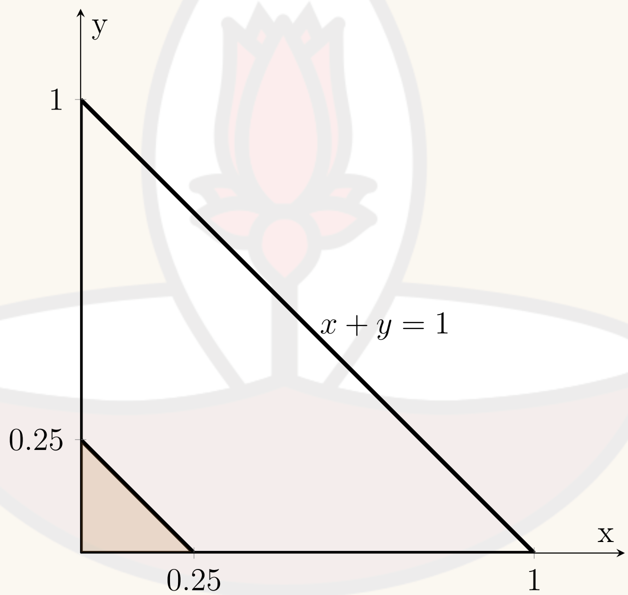
(a) $P(X+Y \le \frac{1}{4}) = \frac{1}{2}$

(b) $P(X+Y \le \frac{1}{2}) = \frac{1}{16}$

(c) X and Y are independent random variables.

(d) X and Y are dependent random variables.

Solution:
Option (a)



Orange region will denote $X+Y \le \frac{1}{4}$. Now,

$$P(X+Y \le \frac{1}{4}) = \int_{y=0}^{1/4}\int_{x=0}^{1/4-y} f_{XY}(x,y)dxdy$$

$$= \int_{y=0}^{1/4}\int_{x=0}^{1/4-y} 24xydxdy$$

$$= \int_{y=0}^{1/4} 12x^2y\Big|_{x=0}^{1/4-y} dy$$

$$= \int_{y=0}^{1/4} 12y \left(\frac{1}{4} - y\right)^2 dy$$

$$= \int_{y=0}^{1/4} \frac{12}{16} y(1-4y)^2 dy$$

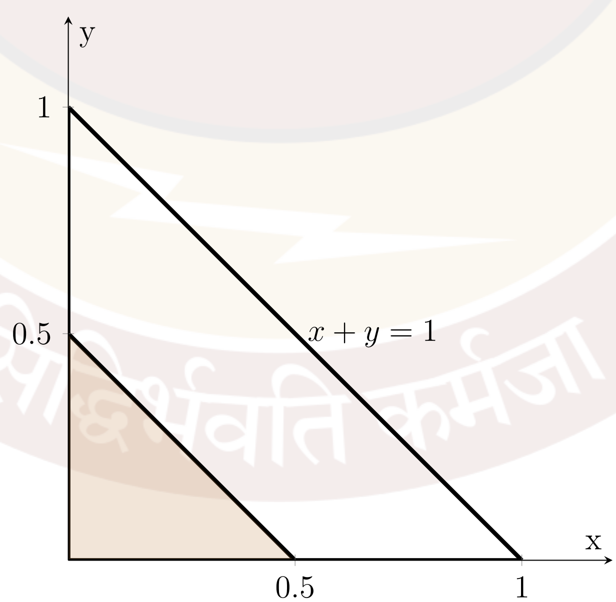
$$= \frac{3}{4} \int_{y=0}^{1/4} y(1+16y^2-8y) dy$$

$$= \frac{3}{4} \left(\frac{y^2}{2} + 4y^4 - \frac{8y^3}{3}\right)\Bigg|_{y=0}^{1/4}$$

$$= \frac{3}{4} \left(\frac{1}{32} + \frac{1}{64} - \frac{1}{24}\right)$$
$$= \frac{3}{4} . \frac{1}{192} = \frac{1}{256}$$

Hence, option (a) is wrong.

Option (b)

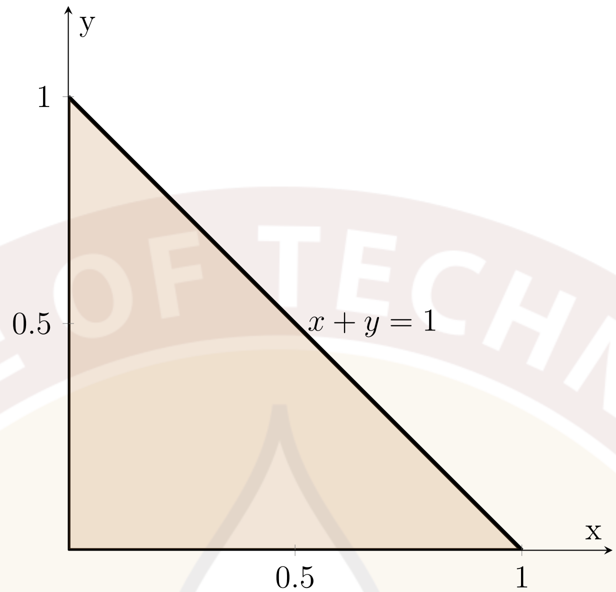


Orange region will denote $X + Y \leq \frac{1}{2}$. Now,

$$
\begin{aligned}
P(X+Y \leq \frac{1}{2}) &= \int_{y=0}^{1/2} \int_{x=0}^{1/2-y} f_{XY}(x,y)dxdy \\
&= \int_{y=0}^{1/2} \int_{x=0}^{1/2-y} 24xydxdy \\
&= \int_{y=0}^{1/2} 12x^2y\Big|_{x=0}^{1/2-y} dy \\
&= \int_{y=0}^{1/2} 12y\left(\frac{1}{2}-y\right)^2 dy \\
&= \int_{y=0}^{1/2} \frac{12}{4} y(1-2y)^2 dy \\
&= 3\int_{y=0}^{1/2} y(1+4y^2-4y)dy \\
&= 3\left(\frac{y^2}{2}+y^4-\frac{4y^3}{3}\right)\Big|_{y=0}^{1/2} \\
&= 3\left(\frac{1}{8}+\frac{1}{16}-\frac{1}{6}\right) \\
&= 3 \times \frac{2}{96} = \frac{1}{16}
\end{aligned}
$$

Hence, option (b) is correct.

<u>Option (c) and (d)</u>



For $0 < x < 1$

$$
\begin{aligned}
f_X(x) &= \int_{y=0}^{1-x} f_{XY}(x,y)dy \\
&= \int_{y=0}^{1-x} 24xydy \\
&= 12xy^2\Big|_{y=0}^{1-x} \\
&= 12x(1-x)^2
\end{aligned}
$$

For $0 < y < 1$

$$
\begin{aligned}
f_Y(y) &= \int_{x=0}^{1-y} f_{XY}(x,y)dx \\
&= \int_{0}^{1-y} 24xydx \\
&= 12x^2y\Big|_{x=0}^{1-y} \\
&= 12y(1-y)^2
\end{aligned}
$$

Therefore, $f_X(x).f_Y(y) = 144xy(1-x)^2(1-y)^2 \neq f_{XY}(x,y)$

Hence, X and Y are not independent.

---

7. The joint pdf of two random variables X and Y is given by

$$f_{XY}(x,y) = \begin{cases} 3xy(1-x) & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Calculate $P(X > \frac{1}{2}|Y = 1)$.
Solution:
We know that

$$P(a < X < b|Y = y) = \frac{f_{XY}(a < X < b, y)}{f_Y(y)}$$

Now,

$$\begin{aligned} f_Y(y) &= \int_0^1 3xy(1-x)dx \\ &= \int_0^1 (3xy - 3x^2y)dx \\ &= \left(\frac{3x^2y}{2} - x^3y\right)\Bigg|_0^1 \\ &= \frac{3y}{2} - y = \frac{y}{2} \end{aligned}$$

Therefore, $f_Y(1) = \frac{1}{2}$
Now,

$$\begin{aligned} P(X > \frac{1}{2}|Y = 1) &= \frac{f_{XY}(X > \frac{1}{2}, Y = 1)}{f_Y(1)} \\ &= 2f_{XY}(X > \frac{1}{2}, Y = 1) \\ &= \int_{x=\frac{1}{2}}^1 2(3x(1-x))dx \\ &= 6\int_{\frac{1}{2}}^1 (x - x^2)dx \\ &= 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Bigg|_{\frac{1}{2}}^1 \\ &= 6\left(\frac{1}{2} - \frac{1}{3}\right) - 6\left(\frac{1}{8} - \frac{1}{24}\right) = 1 - \frac{1}{2} = \frac{1}{2} \end{aligned}$$

---

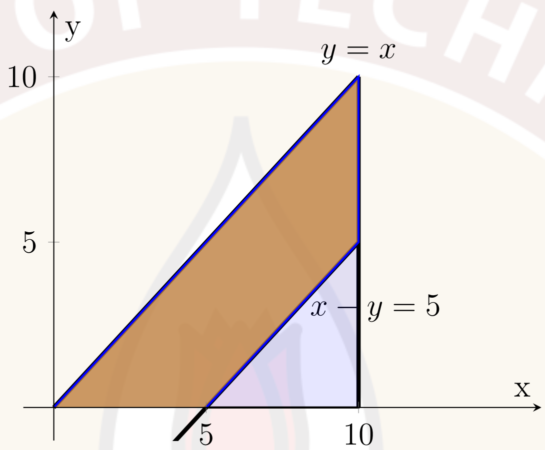
8. The amount of milk (in litres) in a shop at the beginning of any day is a random amount X from which a random amount Y (in litres) is sold during that day. Assume that the

joint density function of X and Y is given by

$$f_{XY}(x,y) = \begin{cases} \frac{1}{50} & 0 \leq x \leq 10, 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that amount of milk left at the end of day is less than 5 litres. Write your answer correct to two decimal points.

Solution:



X denotes the amount of milk at the beginning of any day and Y denotes the amount of milk which is sold during that day.
Therefore, amount of milk left at the end of the day will be denoted by $X - Y$.

To find: $P(X - Y < 5)$

In the diagram above, brown region denotes $X - Y < 5$ and brown + blue region denotes the support of X and Y.

Area of the support$(X, Y) = \frac{1}{2} \times 10 \times 10 = 50$.

Area of brown region = Area of support$(X, Y)-$ area of blue region

$\Rightarrow$ area of brown region $= 50 - \frac{1}{2} \times 5 \times 5 = \frac{75}{2}$

Therefore,

$$\begin{aligned} P(X - Y < 5) &= \frac{\text{area of brown region}}{\text{area of support}} \\ &= \frac{75/2}{50} \\ &= \frac{75}{100} \end{aligned}$$

9. The joint pdf of two continuous random variables X and Y is given by

$$f_{XY}(x,y) = \begin{cases} ke^{-(x+y)} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of $P(X \geq 5, Y \leq 5)$.

(a) e^{-10}

(b) $(e^{-5} - 1)e^{-5}$

(c) $(1 - e^{-5})e^{-5}$

(d) $(e^{-5} + 1)e^{-5}$

Solution:
We know that

$$\int\int_{\text{Supp}(X,Y)} f_{XY} dxdy = 1$$

Therefore,

$$\int_{y=0}^{\infty}\int_{x=0}^{\infty} (ke^{-(x+y)}) dxdy = 1$$
$$\Rightarrow k \int_{y=0}^{\infty}\int_{x=0}^{\infty} e^{-y}e^{-x} dxdy = 1$$
$$\Rightarrow k \int_{y=0}^{\infty} e^{-y}(-e^{-x})\Big|_0^{\infty} dy = 1$$
$$\Rightarrow k \int_{y=0}^{\infty} e^{-y}(0+1)dy = 1$$
$$\Rightarrow k \int_{y=0}^{\infty} e^{-y}dy = 1$$
$$\Rightarrow k(-e^{-y})\Big|_0^{\infty} = 1$$
$$\Rightarrow k(0+1) = 1$$
$$\Rightarrow k = 1$$

To find: $P(X \geq 5, Y \leq 5)$

Now,

$$
\begin{aligned}
P(X \geq 5, Y \leq 5) &= \int_{y=0}^{5} \int_{x=5}^{\infty} (e^{-(x+y)}) dx dy \\
&= \int_{y=0}^{5} \int_{x=5}^{\infty} e^{-y} e^{-x} dx dy \\
&= \int_{y=0}^{5} e^{-y} (-e^{-x}) \Big|_{5}^{\infty} dy \\
&= \int_{y=0}^{5} e^{-y} (0 + e^{-5}) dy \\
&= (e^{-5}) \int_{y=0}^{5} e^{-y} dy \\
&= (e^{-5})(-e^{-y}) \Big|_{0}^{5} \\
&= (e^{-5})(-e^{-5} + 1) \\
&= (e^{-5})(1 - e^{-5})
\end{aligned}
$$

---

10. The joint pdf of two random variables X and Y is given by

$$
f_{XY}(x, y) = \begin{cases} \frac{1}{8}(x+y) & 0 \leq x \leq 2, 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}
$$

Find the value of $P\left(\frac{1}{2} \leq y \leq 1 \mid (X = \frac{1}{2})\right)$. Write your answer correct to two decimal points.

Solution:

We know that

$$
P(a < Y < b | X = x) = \frac{f_{XY}(X = x, a < Y < b)}{f_X(x)}
$$

Now,

$$
\begin{aligned}
f_X(x) &= \int_0^2 \frac{1}{8}(x+y) dy \\
&= \frac{1}{8}\left(xy + \frac{y^2}{2}\right)\Big|_0^2 \\
&= \frac{2x+2}{8} = \frac{x+1}{4}
\end{aligned}
$$

Therefore, $f_X(\frac{1}{2}) = \frac{3}{8}$

Now,

$$P(\frac{1}{2} \leq Y \leq 1 | X = \frac{1}{2}) = \frac{f_{XY}(X = \frac{1}{2}, \frac{1}{2} \leq Y \leq 1)}{f_X(\frac{1}{2})}$$

$$= \int_{1/2}^{1} \frac{8}{3} \left[\frac{1}{8} \left(\frac{1}{2} + y \right) \right] dy$$

$$= \int_{1/2}^{1} \frac{1}{3} \left(\frac{1}{2} + y \right) dy$$

$$= \left(\frac{y}{6} + \frac{y^2}{6} \right) \Bigg|_{1/2}^{1}$$

$$= \left(\frac{1}{6} + \frac{1}{6} \right) - \left(\frac{1}{12} + \frac{1}{24} \right)$$

$$= \frac{1}{3} - \frac{1}{8} = \frac{5}{24} = 0.20$$
