Homework 2 - Probability

Thiều Ngọc Mai

August 2023

Problem 1

a. The marginal distributions:

$$
\begin{aligned}
&P(X = x_1) = 0.01 + 0.05 + 0.1 = 0.16\\
&P(X = x_2) = 0.02 + 0.1 + 0.05 = 0.17\\
&P(X = x_3) = 0.03 + 0.05 + 0.03 = 0.11\\
&P(X = x_4) = 0.1 + 0.07 + 0.05 = 0.22\\
&P(X = x_5) = 0.1 + 0.2 + 0.04 = 0.34\\
&P(Y = y_1) = 0.01 + 0.02 + 0.03 + 0.1 + 0.1 = 0.26\\
&P(Y = y_2) = 0.05 + 0.1 + 0.05 + 0.07 + 0.2 = 0.47\\
&P(Y = y_3) = 0.1 + 0.05 + 0.03 + 0.05 + 0.04 = 0.27
\end{aligned}
$$

b. The conditional distributions:

- Have:

  $$P(X = x_1|Y = y_1) = \frac{P(X = x_1, Y = y_1)}{P(Y = y_1)} = \frac{0.01}{0.26} = \frac{1}{26}(*1)$$

  Use the similar formula for (*1), substitute $X =$

x_2, x_3, x_4, x_5 to (*1):

$$P(X = x_2|Y = y_1) = \frac{1}{13}$$
$$P(X = x_3|Y = y_1) = \frac{3}{26}$$
$$P(X = x_4|Y = y_1) = \frac{5}{13}$$
$$P(X = x_5|Y = y_1) = \frac{5}{13}$$

- Have:

  $$P(Y = y_1|X = x_3) = \frac{P(Y = y_1, X = x_3)}{P(X = x_3)} = \frac{0.03}{0.11} = \frac{3}{11}(*2)$$

  Use the similar formula for (*2), substitute $Y = y_2, y_3$ to (*2):

  $$P(Y = y_2|X = x_3) = \frac{5}{11}$$
  $$P(Y = y_3|X = x_3) = \frac{3}{11}$$

Problem 2

- For the discrete random variables:

  The expected value of discrete random variables is defined as:

  $$E[X] = \sum_x xp(x)$$

The conditional expectation of discrete random variables is defines as:

$$E_x[x|y] = \sum_x xp(x|y)$$

We have:

$$\begin{aligned}
E_y[E_x[x|y]] &= \sum_y E_x[x|y]p(y) \\
&= \sum_y \sum_x xp(x|y)p(y) \\
&= \sum_y \sum_x xp(x,y) \\
&= \sum_x x \sum_y p(x,y) \\
&= \sum_x xp(x) \\
&= E[x]
\end{aligned}$$

- For the continuous random variables:
  The expected value of continuous random variables is defined as:

  $$E[X] = \int\limits_{-\infty}^{\infty} xf(x)dx$$

  The conditional expectation of continuous random variables is defines as:

  $$E_x[x|y] = \int\limits_{-\infty}^{\infty} xf_{X|Y}(x|y)dx$$

We have:

$$
\begin{aligned}
E_y[E_x[x|y]] &= \int\limits_{-\infty}^{\infty} E_x[x|y]f(y)dy \\
&= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x f_{X|Y}(x|y)dx f(y)dy \\
&= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x f_{XY}(x,y)dxdy \\
&= \int\limits_{-\infty}^{\infty} x \int\limits_{-\infty}^{\infty} f_{XY}(x,y)dxdy \\
&= \int\limits_{-\infty}^{\infty} x f_X(x)dx \\
&= E[x]
\end{aligned}
$$

Problem 3

We have:

$$\begin{aligned}
V_x &= E[(X-\mu)^2] \\
&= \sum_x (x-\mu)^2 p(x) \\
&= \sum_x (X^2 - 2x\mu + \mu^2) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= E[x^2] - 2\mu^2 + \mu^2 \\
&= E[x^2] - \mu^2 \\
&= E[x^2] - (E[x])^2
\end{aligned}$$

Problem 4

Events:
C = Tumor is cancerous
B = Tumor is benign

$P(C) = 0.01 \Rightarrow$P(B) $= 1 - 0.01 = 0.99$
$P(pos|C) = 0.8$
$P(neg|B) = 0.9 \Rightarrow$P(pos|B) $= 0.1$

Apply Baye's Formula:

$$\begin{aligned} P(C|pos) &= \frac{P(pos|C)P(C)}{P(pos)} \\ &= \frac{P(pos|C)P(C)}{P(pos|C)P(C) + P(pos|B)P(B)} \\ &= \frac{0.8 \times 0.01}{0.8 \times 0.01 + 0.1 \times 0.99} \\ &= 0.075 \end{aligned}$$

Thus, the estimation of physicians is inaccurate.

Problem 5

- Event:
  C = The door you pick has pure gold car
  E = The evidence the host has revealed a door with gummy bear.

- According to the Baye's Theorem :

  $$\begin{aligned} P(C|E) &= \frac{P(E|C)P(C)}{P(E)} \\ &= \frac{P(E|C)P(C)}{P(E|C)P(C) + P(E|C^c)P(C^c)} \end{aligned}$$

  At the beginning, you have no information about what behind the door, thus $P(C) = \frac{1}{4} \Rightarrow \mathrm{P(C^c)} = \frac{3}{4}$

  $P(E|C)$ is the probability the host would show the door with gummy bear, given that the door you chose has the car behind. Since the host always opens the door with gummy bear, $P(E|C) = 1$

$P(E|C^c)$ is the probability the host would show the door with gummy bear, given that the door you chose doesn't have the car behind. Since the host always opens the door with gummy bear, $P(E|C^c) = 1$

Then,

$$P(C|E) = \frac{1 \times \frac{3}{4}}{1 \times \frac{1}{4} + 1 \times \frac{3}{4}} = \frac{1}{4}$$

- After reveal one door, the probability the car is behind the door you chose is unchanged. The probability the car is in the 2 others door now becomes $\frac{3}{4}$. However, since we have no information about that 2 doors, the probability the car is in one of that 2 doors is $\frac{1}{2}$.

  Therefore, if you switch door, the probability to win is $\frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$
- $\frac{a}{b} = \frac{3}{8} \Rightarrow a + b = 3 + 8 = 11$