Statistics for Data Science - 2

Week 2 practice Assignment

Discrete random variable

1. A random variable X is defined as the length of the hypotenuse of the right-angled triangle whose other two sides are determined by the roll of two 6-sided dice. How many values does X take? [1 mark]

   Solution:
   When two dice are rolled then there are a total of 36 outcomes.
   The outcomes are:
   {(1, 1) , (1, 2), ... , (1, 6),
   (2, 1), (2, 2), ... , (2, 6),
   ...
   ...
   (6, 1), (6, 2), ... , (6, 6)}

   But the outcomes like (1, 2) (2, 1) will give the same length of the hypotenuse, hence a total of 21 values are possible for the random variable X.

---

2. Two cards are drawn from a well shuffled pack of 52 cards one after other without replacement. A random variable is defined as:

$$X = \begin{cases} 0 & \text{if both cards are of same color} \\ 1 & \text{if both cards are of different color} \end{cases}$$

   Find the probability mass function of X. [1 mark]

   (a)

x	0	1
$f_X(x)$	$\frac{1}{13}$	$\frac{12}{13}$

   (b)

x	0	1
$f_X(x)$	$\frac{1}{2}$	$\frac{1}{2}$

   (c)

x	0	1
$f_X(x)$	$\frac{25}{51}$	$\frac{26}{51}$

   (d)

x	0	1
$f_X(x)$	$\frac{12}{25}$	$\frac{13}{25}$

   Solution:

$$P(X=0) = P(\text{Both the cards are of same colors})$$
$$= P(\text{First card is any one of 52 cards}).P(\text{2nd card is of same color as of 1st card})$$
$$= 1.\frac{25}{51}$$

$$P(X=1) = P(\text{Both the cards are of different colors})$$
$$= P(\text{First card is any one of 52 cards}).P(\text{2nd card is of different color as of 1st card})$$
$$= 1.\frac{26}{51}$$

Hence, option (c) is right.

---

3. In a group of fifteen people, 8 people have blood group type O, 4 people have blood group type A, and 3 people have blood group type B. If five people are selected randomly from these fifteen people, then what is the probability that out of these five people 2 people have blood group type O, 2 have blood group type A and one has blood group type B? (Answer the question correct up to two decimal places.) [2 mark]

Solution:
Number of ways of selecting five people out of 15 = ${}^{15}C_5$
Number of ways of selecting 2 people of blood group of type O out of 8 people of blood group of type O= ${}^{8}C_2$
Number of ways of selecting 2 people of blood group of type A out of 4 people of blood group of type A= ${}^{4}C_2$
Number of ways of selecting 1 people of blood group of type B out of 3 people of blood group of type B= ${}^{3}C_1$
Therefore, required probability = $\dfrac{{}^{8}C_2 {}^{4}C_2 {}^{3}C_1}{{}^{15}C_5}$
. $= \dfrac{28 \times 6 \times 3}{3003} = 0.167$

---

4. Probability mass funcion of a discrete random variable X is given as:

x	-2	-1	0	1	2
$f_X(x)$	a	0.2	b	0.1	0.2

Table 2.1.P: PMF of X

If $P(X \leq 1 | X \geq -1) = \dfrac{3}{4}$, then find the value of $P(X = -2)$. [2 marks]
Solution:

We know that

$$\sum_{x \in T_X} f_X(x) = 1$$
$$\Rightarrow a + 0.2 + b + 0.1 + 0.2 = 1$$
$$\Rightarrow a + b = 0.5 \qquad ...(1)$$

From the given condition, we have

$$P(X \leq 1 | X \geq -1) = \frac{3}{4}$$
$$\Rightarrow \frac{P(X \leq 1, X \geq -1)}{P(X \geq -1)} = \frac{3}{4}$$
$$\Rightarrow \frac{P(\{-1, 0, 1\})}{P(\{-1, 0, 1, 2\})} = \frac{3}{4}$$
$$\Rightarrow \frac{b + 0.3}{b + 0.5} = \frac{3}{4}$$
$$\Rightarrow 4b + 1.2 = 3b + 1.5$$
$$\Rightarrow b = 0.3 \qquad ...(2)$$

From equations (1) and (2), we have

$$a = 0.2$$
$$b = 0.3$$

$$P(X = -2) = a$$
$$\Rightarrow P(X = -2) = 0.2$$

---

5. Siberian seagulls migrate to Ganga river to escape harsh winter weather in the months of October to March. It is seen that the number of Siberian seagulls reaching Ganga river on one day in January is Poisson distributed with an average of 1000. What is the probability that 650 seagulls will arrive on a given day of January? [2 marks]

   (a) $\frac{e^{-650}(650)^{1000}}{650!}$

   (b) $\frac{e^{-650}(650)^{1000}}{1000!}$

   (c) $\frac{e^{-1000}(650)^{1000}}{650!}$

(d) $\dfrac{e^{-1000}(1000)^{650}}{650!}$

Solution:
Let X be the number of Siberian seagulls migrating everyday near to Ganga river. By given condition, we have

$$X \sim \text{Poisson}(1000)$$

$$\begin{aligned} P(X = x) &= \frac{e^{-\lambda}\lambda^x}{x!} \\ \Rightarrow P(X = 650) &= \frac{e^{-1000}(1000)^{650}}{650!} \end{aligned}$$

---

6. Probability mass function of a discrete random variable X is given as:

x	-1	0	1	2	3
$f_X(x)$	0.1	0.3	0.2	0.1	0.3

Table 2.2.P: PMF of X

If another random variable Y is defined as $Y = X(X-1)$, then find the smallest value of y in the range of Y such that $P(Y \leq y) > \frac{1}{2}$ and $P(Y \geq y) \leq \frac{1}{2}$. [2 marks]

Solution:
Y is defined as $Y = X(X-1)$

At $X = -1, Y = -1(-2) = 2$
At $X = 0, Y = 0(-1) = 0$
At $X = 1, Y = 1(0) = 0$
At $X = 2, Y = 2(1) = 2$
At $X = 3, Y = 3(2) = 6$

Therefore, $T_Y = \{0, 2, 6\}$

$P(Y = 0) = P(X \in \{0, 1\}) = 0.3 + 0.2 = 0.5$
$P(Y = 2) = P(X \in \{-1, 2\}) = 0.1 + 0.1 = 0.2$
$P(Y = 6) = P(X = 3) = 0.3$

Now,

$$P(Y \leq 0) = P(Y = 0) = 0.5$$

First required condition is not satisfied at $Y = 0$.

$$P(Y \leq 2) = P(Y = 0) + P(Y = 2) = 0.5 + 0.2 = 0.7$$

Both the required conditions are satisfied at $Y = 2$.

---

7. Three friends toss three fair coins to decide who is going to pay for the dinner. The person getting an outcome different from the other two outcomes will pay for the dinner. If all three coins result in the same outcome, they will toss the coins again. If X denotes the number of trials needed to decide who is going to pay, then what is the probability that X is at most 3? (Answer the question correct up to two decimal places.) [2 marks]
Solution:
Let X be the number of trials to decide who is going to pay.
Sample space on tossing three coins are:
{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT }
P(They will decide who is going to pay) $= P(\{$ HHT, HTH, THH, HTT, THT, TTH $\}) = \frac{6}{8} = \frac{3}{4}$
P(They will not decide who is going to pay) $= P(\{$ HHH, TTT $\}) = \frac{2}{8} = \frac{1}{4}$
X will take values as 1, 2, 3, 4, ...
and $X \sim$ Geometric($\frac{3}{4}$)

$$\begin{aligned} P(X \leq 3) &= P(X = 1) + P(X = 2) + P(X = 3) \\ &= \frac{3}{4} + \frac{1}{4}.\frac{3}{4} + \left(\frac{1}{4}\right)^2.\frac{3}{4} \\ &= 0.98 \end{aligned}$$

---

8. Let $X \sim$ Uniform($\{1, 2, 3, ... \ n\}$). If the probability that X is an odd number is $\frac{6}{11}$, then what can be the value of n? [2 marks]

   (a) 11 only

   (b) 12 only

   (c) Any multiple of 11.

   (d) Any odd multiple of 11.

Solution:
Since, $X \sim$ Uniform($\{1, 2, 3, ... \ n\}$)
Let A be the event that X takes odd numbers.
Therefore,

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} \qquad ...(1)$$

where $S = \{1, 2, 3, ...n\}$

It is given that

$$P(A) = \frac{6}{11} \qquad ...(2)$$

By equation (1) and (2), we have
n should be multiple of 11 and number of odd numbers less than or equal to n should be multiple of 6.
This is possible only for $n = 11$.

---

9. The number of customers arriving per day at a certain automobile service facility is assumed to follow a Poisson distribution with an average of 50 customers arriving each day. Assume that number of customers on different days are independent. What is the probability that exactly 40 customers will come for at least 5 days over a 30 days period? [3 marks]

   (a) $1 - \sum_{x=0}^{4} \left({}^{30}C_x \left(\frac{e^{-50}(50)^{40}}{40!} \right)^x \left(1 - \frac{e^{-50}(50)^{40}}{40!} \right)^{30-x} \right)$

   (b) $\sum_{x=0}^{4} \left({}^{30}C_x \left(\frac{e^{-50}(50)^{40}}{40!} \right)^x \left(1 - \frac{e^{-50}(50)^{40}}{40!} \right)^{30-x} \right)$

   (c) ${}^{30}C_5 \left(\frac{e^{-50}(50)^{40}}{40!} \right)^5 . \left(1 - \frac{e^{-50}(50)^{40}}{40!} \right)^{25}$

   (d) ${}^{30}C_5 \left(1 - \frac{e^{-50}(50)^{40}}{40!} \right)^5 . \left(\frac{e^{-50}(50)^{40}}{40!} \right)^{25}$

   Solution:
   Let X be the number of customers arriving per day at a certain automobile service facility.
   $X \sim Poisson(50)$

   $$P(X = 40) = \frac{e^{-50}50^{40}}{40!}$$

   Let Y be the number of days in the next 30 days on which 40 customers have arrived on that particular shop.
   Then, $Y \sim \text{Binomial}\left(30, \frac{e^{-50}50^{40}}{40!}\right)$
   Now,

   $$P(Y \geq 5) = 1 - P(Y < 5)$$

   $$1 - \sum_{x=0}^{4} \left({}^{30}C_x \left(\frac{e^{-50}(50)^{40}}{40!} \right)^x \left(1 - \frac{e^{-50}(50)^{40}}{40!} \right)^{30-x} \right)$$

---

10. A biased coin with the probability of 0.4 of showing head is tossed until it shows either two consecutive heads or two consecutive tails. If X denotes the number of tosses required, what is the value of $P(X = 5)$? [3 marks]

   (a) 0.03456
   (b) 0.02304
   (c) 0.01675
   (d) 0.0576

   Solution:
   It is clear that

$$\begin{aligned} P(X = 5) &= P(\mathrm{HTHTT}) + P(\mathrm{THTHH}) \\ &= (0.4)^2(0.6)^3 + (0.4)^3(0.6)^2 \\ &= 0.0576 \end{aligned}$$