Statistics for Data Science - 2

Week 1 Practice Assignment Solution

Events and probabilities

1. A customer will purchase a shirt with probability 0.5. The customer will purchase a pant with probability 0.4 and will purchase both a shirt and a pant with probability 0.2. What is the probability that the customer will purchase neither a shirt nor a pant?

   Solution:

   Let A be the event that the customer will purchase a shirt and B be the event that the customer will purchase a pant.
   Given that, $P(A) = 0.5$ and $P(B) = 0.4$.
   Also given that the customer will purchase both a shirt and a pant with probability 0.2. i.e. $P(A \cap B) = 0.2$.
   We have to find the probability that the customer will purchase neither a shirt nor a pant i.e. $P(A^C \cap B^C)$.
   We know that $P(A^C \cap B^C) = P((A \cup B)^C) = 1 - P(A \cup B)$
   And, $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.4 - 0.2 = 0.7$
   $\Rightarrow P(A^C \cap B^C) = 1 - P(A \cup B) = 1 - 0.7 = 0.3$

2. Suppose that we roll a pair of fair dice, so each of the 36 possible outcomes is equally likely. Let A denote the event that the first die shows 5, B be the event such that the sum of the outcomes of rolling the pair of dice is 10, and C be the event such that the sum of the outcomes of rolling the pair of dice is 7. Then

   a) Event A and event B are independent.

   b) Event A and event B are not independent.

   c) Event A and event C are independent.

   d) Event A and event C are not independent.

   Solution:

   We are rolling a pair of fair dice and all the 36 outcomes is equally likely that means probability of occurring each outcome is same i.e. $1/36$.
   A is the event that the first die shows 5.
   $\Rightarrow A = \{(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)\}$
   B is the event that the sum of the outcomes of rolling the pair of dice is 10.
   $\Rightarrow B = \{(4,6),(5,5),(6,4)\}$
   C is the event that the sum of the outcomes of rolling the pair of dice is 7.
   $\Rightarrow C = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$
   Also, $A \cap B = \{(5,5)\}$ and $A \cap C = \{(5,2)\}$
   Since each outcome is equally likely, so

$P(A) = \frac{6}{36}, P(B) = \frac{3}{36}, P(C) = \frac{6}{36}, P(A \cap B) = \frac{1}{36}$ and $P(A \cap C) = \frac{1}{36}$

Since $P(A \cap B) = \frac{1}{36} \neq P(A)P(B) \Rightarrow$ event A and B are not independent.

Also, $P(A \cap C) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = P(A)P(C) \Rightarrow$ event A and C are independent.
Hence, option (b) and (c) are correct.

3. Let A and B be two independent events of a random experiment. Then, which of the following is/are always true?

   a) $P(A \cup B) = P(A)P(B) + P(B)$
   b) $P(A \cup B) = P(A)P(B^C) + P(B)$
   c) $P(A \cup B) = P(A) + P(B)$
   d) $P((A \cap B)|A) = P(B)$

   Solution:
   Given that A and B are two independent events $\Rightarrow P(A \cap B) = P(A)P(B)$.

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A)P(B) \\ &= P(A)[1 - P(B)] + P(B) \\ &= P(A)P(B^C) + P(B) \end{aligned}$$

   Therefore, option (b) is correct.
   Consider

$$\begin{aligned} P((A \cap B)|A) &= \frac{P((A \cap B) \cap A)}{P(A)} \\ &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(A)P(B)}{P(A)} \\ &= P(B) \end{aligned}$$

   This implies that option (d) is also correct.
   Hence, option (b) and (d) are correct.

4. The probability that a student registered for IITM online degree program will pass the qualifier exam is 0.6 independent of all other students. Find the probability that out of 10,000 registered students, 7,000 students will pass the qualifier exam.

a) $(0.6)^{3000}(0.4)^{7000}$

b) $(0.6)^{7,000}(0.4)^{3,000}$

c) ${}^{10,000}C_{7,000}(0.6)^{3,000}(0.4)^{7,000}$

d) ${}^{10,000}C_{7,000}(0.6)^{7,000}(0.4)^{3,000}$

Solution:
Probability(p) that the student registered for IITM online degree program will pass the qualifier exam is 0.6.
We have to find the probability that out of 10,000 registered students, 7,000 students will pass the qualifier exam and passing qualifier exam for any student will be independent of the other.
So here we can use binomial distribution with X will be number of students who will pass the exam along with $p = 0.6, n = 10,000$, and $k = 7,000$.
And we know that for binomial distribution $P(X = k) = {}^nC_k p^k(1-p)^{(n-k)}$

$$\Rightarrow P(X = 7,000) = {}^{10,000}C_{7,000}(0.6)^{7,000}(1-0.6)^{(10,000-7,000)}$$

$$\Rightarrow P(X = 7,000) = {}^{10,000}C_{7,000}(0.6)^{7,000}(0.4)^{3,000}$$

Hence, probability that out of 10,000 registered students, 7,000 students will pass the qualifier exam is ${}^{10,000}C_{7,000}(0.6)^{7,000}(0.4)^{3,000}$.

5. Assume that the probability of a defective computer component is 0.05. Components are randomly selected for being tested(assume that the testing is 100% accurate). Find the probability that the first defect is observed when the sixth component is tested.

   a) $(0.05)^6 \times 0.95$

   b) $(0.95)^6 \times 0.05$

   c) $(0.95)^5 \times 0.05$

   d) $(0.05)^5 \times 0.95$

Solution:
We have to find the probability that the first defect is observed when the sixth component is tested.
The probability of a defective computer component is 0.05.
Here we can assume that getting a defective component is success. That means we have to find the probability of first success at 6th trials with p given as 0.05.
So here we can use geometric distribution with X representing the number of components tested along with $p = 0.05$
And we know that for geometric distribution $P(X = k) = (1-p)^{k-1}p$.

$\Rightarrow P(X = 6) = (1 - 0.05)^{6-1} \times 0.05$

$\Rightarrow P(X = 6) = (0.95)^5 \times 0.05$

Hence the probability that the first defect is observed when the sixth component is tested is $(0.95)^5 \times 0.05$.

6. If Aarushi and Ansh play a game of chess, Aarushi wins with probability 0.5 and Ansh wins with probability 0.4 and the game ends in a draw with probability 0.1, independent of all other games. They agree to play a match consisting of 5 games. Find the probability that Aarushi wins 4-1 (win gives 1 pt to winner and draw gives 0.5 pts to both). Enter your answer correct to 3 decimals accuracy.
   Solution:
   Let A_i be the event that Aarushi will win the ith game and B_j be the event that Ansh will win the jth game.
   From given information we have $P(A_i) = 0.5, P(B_j) = 0.4$
   There are two disjoint ways that Aarushi wins 4-1.
   i) Aarushi wins 4 games and Ansh wins one game.
   Probability of happening this will be ${}^5C_4(0.5)^4 \times 0.4 = 0.125$
   ii) Aarushi wins 3 games and 2 games are drawn.
   Probability of happening this will be ${}^5C_3(0.5)^3 \times (0.1)^2 = 0.0125$
   So, the probability that Aarushi wins 4-1 is 0.125 + 0.0125 = 0.1375

7. The probability of someone catching flu in a particular winter when they have been given the flu vaccine is 0.2. Without the vaccine, the probability of catching flu is 0.5. If 40% of the population has been given the vaccine, what is the probability that a person chosen at random from the population will catch flu over that winter? Enter the answer correct to 2 decimals accuracy.
   Solution:
   Let A be the event that the person will catch flu and V be the event that the person has been given the vaccine.
   Given that $P(A \mid V) = 0.2, P(A \mid V^C) = 0.5$ and $P(V) = 0.4$
   We have to find the probability that a person chosen at random from the population will catch flu over that winter i.e. $P(A)$.
   And we can write $P(A) = P(A \mid V)P(V) + P(A \mid V^C)P(V^C)$
   $\Rightarrow P(A) = 0.2 \times 0.4 + 0.5 \times (1 - 0.4)$
   $\Rightarrow P(A) = 0.38$

8. Suppose you are playing a game of cards with your friend. Your friend is supposed to give you 13 cards one by one. With a well-shuffled pack of 52 cards, what is the probability that you are dealt a perfect hand(13 of one suit)?

   a) $\dfrac{13!}{52!}$

b) $\boxed{\dfrac{12! \times 39!}{51!}}$

c) $\dfrac{13! \times 39!}{51!}$

d) $\dfrac{13! \times 39!}{52!}$

Solution:
Your friend is supposed to give you 13 cards one by one. Need to find the probability that you are dealt a perfect hand i.e. you have gotten 13 cards of one suit.
For the first card, it can be any card from the 52 cards so probability will be 1.
Once the first card is given to you, the probability for the second card to be of same suit will be $\frac{12}{51}$ because once the first card is given to you it will belong to one particular suit and second card will be conditional on that.
Similarly for the third card, probability will be $\frac{11}{50}$.
Continue like this, we get that the probability that you are dealt a perfect hand is

$$= 1 \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48} \times \frac{8}{47} \times \frac{7}{46} \times \frac{6}{45} \times \frac{5}{44} \times \frac{4}{43} \times \frac{3}{42} \times \frac{2}{41} \times \frac{1}{40}$$
$$= \frac{12! \times 39!}{51!}$$

9. A person has bought a bed from an online furniture store. The seller delivers the disassembled bed parts along with some screws to assemble it. The probability of a screw being defective is 0.1 independent of all other screws. To compensate for the manufacturing error, the seller sends two extra screws in the package where the bed needs exactly 8 screws to assemble. What is the probability that the buyer will be able to assemble the bed? (Enter the answer correct to 4 decimal accuracy)
Solution:
Let X represents the number of screws that seller sends with the bed.
We need exactly 8 screws to assemble the bed and the seller sends two extra i.e. seller sends ten screws.
The buyer will be able to assemble the bed if 8 screws are non - defective or 9 screws are non - defective or 10 screws are non - defective out of the ten screws.
We can relate this with binomial distribution as $X \sim Binomial(10, p)$ where p is the probability of a screw being non - defective and value of p will be 1 - 0.1 = 0.9
The buyer will be able to assemble the bed if at least 8 screws are non - defective.
So, the probability that the buyer will be able to assemble the bed is $P(X \geq 8)$.

And

$$\begin{aligned}
P(X \geq 8) &= P(X = 8) + P(X = 9) + P(X = 10) \\
&= {}^{10}C_8(0.9)^8(0.1)^2 + {}^{10}C_9(0.9)^9(0.1)^1 + {}^{10}C_{10}(0.9)^{10}(0.1)^0 \\
&= (0.9)^8[(0.1)^2 \times 45 + 10 \times 0.9 \times 0.1 + 0.81] \\
&= (0.9)^8 \times 2.16 \\
&= 0.9298
\end{aligned}$$

10. In a pizza shop 40% of the customers order medium size pizza, 50% order small size pizza, and 10% order large size pizza. Of those ordering medium size pizza $\frac{2}{3}$ also ask to add extra toppings. Of those ordering small size pizza $\frac{1}{5}$ also ask to add extra toppings, and of those ordering large size pizza $\frac{4}{5}$ also ask to add extra toppings. Given that a customer asked to add extra toppings, find the conditional probability that the customer ordered a medium pizza.

    a) $\frac{15}{67}$
    b) $\frac{40}{67}$
    c) $\frac{12}{67}$
    d) $\frac{52}{67}$

    Solution:
    Let S, M and L denote the event that customer will order small, medium and large size pizza, respectively.
    Given that $P(S) = 0.50, P(M) = 0.40$ and $P(L) = 0.10$.
    Also, let T be the event that customer will ask to add extra toppings.
    This implies that $P(T \mid S) = \frac{1}{5}, P(T \mid M) = \frac{2}{3}$ and $P(T \mid L) = \frac{4}{5}$.
    We need to find $P(M \mid T)$.
    And

$$\begin{aligned}
P(M \mid T) &= \frac{P(M \cap T)}{P(T)} \\
&= \frac{P(T \mid M)P(M)}{P(T \mid S)P(S) + P(T \mid M)P(M) + P(T \mid L)P(L)} \\
&= \frac{\frac{2}{3} \times 0.40}{\frac{1}{5} \times 0.50 + \frac{2}{3} \times 0.40 + \frac{4}{5} \times 0.10} \\
&= \frac{0.80}{3} \times \frac{15}{6.7} \\
&= \frac{40}{67}
\end{aligned}$$