Anmol Saraf
Email: 200070007@iitb.ac.in
Course: DS203 - Programming for Data Science Minor
Instructor: Dr. Manjesh K Hanawal

Assignment - 1
Roll No.: 200070007
Semester: Autumn 2021
Due Date: 12^{th} August, 2021

Exercise 1
A product is manufactured by two factories A and B. 80% of the product is manufactured in company A and rest in company B. 30% of the product manufactured by company A are defective while 10% of the product manufactured by company B are defective. A sample of the product is randomly selected from the market. Then,

1. What is the probability that the sample is defective.
2. What is the probability that a defective sample in the market is manufactured at company A.

Solution: Let A and B are events that the product is manufactured at the factory A or B respectively and let D be the event where the product is defective.

Given P(A) = 0.8, P(D|A) = 0.3 and P(D|B) = 0.1.

Part 1 Since A and B are mutually exclusive and $A \cup B = \Omega$,

$$\begin{aligned} P(D) &= P(A)P(D|A) + P(B)P(D|B) \\ &= 0.8 * 0.3 + 0.2 * 0.1 \\ &= 0.26 \end{aligned}$$

Part 2 Since the product is already defective (D has occured), we have to find P(A|D),

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{P(A)P(D|A)}{P(D)}$$
$$P(A|D) = \frac{0.8 * 0.3}{0.26} = \frac{12}{13} \approx 0.92$$

Exercise 2
A particular webserver may be working or not working. If the webserver is not working, any attempt to access it fails. Even if the webserver is working, an attempt to access it can fail due to network congestion beyond the control of the webserver. Suppose that the a priori probability that the server is working is 0.8. Suppose that if the server is working, then each access attempt is successful with probability 0.9, independently of other access attempts. Find the following quantities.

1. P(first access attempt fails)
2. P(server is working | first access attempt fails)
3. P(second access attempt fails | first access attempt fails)
4. P(server is working | first and second access attempts fail).

Solution: Let the event W be such that the webserver is working and S be the event where the access attempt was successful.

Given P(W) = 0.8, P(S|W^c) = 0 and P(S|W) = 0.9.

Part 1

$$\begin{aligned}P(S^c) = 1 - P(S) &= 1 - \{P(W)P(S|W) + P(W^c)P(S|W^c)\} \\ &= 1 - (0.8 * 0.9 + 0.2 * 0) = 0.28\end{aligned}$$

Part 2

$$\begin{aligned}P(W|S^c) = \frac{P(W \cap S^c)}{P(S^c)} = \frac{P(W)P(S^c|W)}{P(S^c)} &= \frac{P(W)\,(1 - P(S|W))}{P(S^c)} \\ &= \frac{0.8 * (1 - 0.9)}{0.28} \\ &= \frac{2}{7} \approx 0.29\end{aligned}$$

Part 3

$$P(S_2^c|S_1^c) = \frac{P(S_2^c \cap S_1^c)}{P(S_1^c)} = \frac{P(W^c)P(S_2^c \cap S_1^c|W^c) + P(W)P(S_2^c \cap S_1^c|W)}{P(S_1^c)}$$

Since we know that P(S|W^c) = 0,

$$P(S_2^c|S_1^c) = \frac{0.2 * 1 + 0.8 * 0.1 * 0.1}{0.28} = \frac{26}{35} \approx 0.74$$

Part 4

$$\begin{aligned}P(W|S_2^c \cap S_1^c) &= \frac{P(W)P(S_2^c \cap S_1^c|W)}{P(W^c)P(S_2^c \cap S_1^c|W^c) + P(W)P(S_2^c \cap S_1^c|W)} \\ &= \frac{0.8 * 0.1 * 0.1}{0.2 * 1 + 0.8 * 0.1 * 0.1} = \frac{1}{26} \approx 0.038\end{aligned}$$

Exercise 3
Two dice are rolled. What is the probability that at least one is a six? If the two faces are different, what is the probability that at least one is a six?

Solution: Let A and B be the events that a 6 is rolled on the die 1 or 2 respectively. Therefore, probability such that at least one is a 6,

$$\begin{aligned}P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{1}{6} + \frac{1}{6} - \frac{1}{6} * \frac{1}{6} = \frac{11}{36} \approx 0.31\end{aligned}$$

Now, let the event when two of the faces are different be E, therefore

$$P(E) = 6 * \frac{1}{6} * \frac{5}{6} = \frac{5}{6}$$

Now, the probabilty that atleast one face is 6 when both are not same is,

$$\begin{aligned}P(A \cup B|E) &= \frac{P((A \cup B) \cap E)}{P(E)} = \frac{P(A \cup B) - P(A \cap B)}{P(E)} \\ &= \frac{\frac{11}{36} - \frac{1}{36}}{\frac{5}{6}} = \frac{\frac{5}{18}}{\frac{5}{6}} = \frac{1}{3} \approx 0.33\end{aligned}$$

Exercise 4
Suppose that 5 percent of men and 1 percent of women are color-blind. A color-blind person is chosen at random. What is the probability of this person being male? Assume that there are an equal number of males and females.

Solution: Let the events M, F and C be the events that a person choosen is male, female or color-blind respectively.
Given P(M) = P(F) = 0.5, P(C|M) = 0.05 and P(C|F) = 0.01.
Therefore, using Total Probability Law to find P(C),

$$\begin{aligned} P(C) &= P(M)P(C|M) + P(F)P(C|F) \\ &= 0.5 * 0.05 + 0.5 * 0.01 = 0.03 \end{aligned}$$

Now, using Bayes' Formula to find P(M—C),

$$\begin{aligned} P(M|C) &= \frac{P(M \cap C)}{P(C)} = \frac{P(M)P(C|M)}{P(C)} \\ &= \frac{0.5 * 0.05}{0.03} = \frac{5}{6} \approx 0.83 \end{aligned}$$

Exercise 5
(a) Suppose that an event E is independent of itself. Show that either P(E) = 0 or P(E) = 1.

(b) Events A and B have probabilities P(A) = 0.3 and P(B) = 0.4. What is P(A∪B) if A and B are independent? What is P(A∪B) if A and B are mutually exclusive?

(c) Now suppose that P(A) = 0.6 and P(B) = 0.8. In this case, could the events A and B be independent? Could they be mutually exclusive?

Solution: Two events A and B are said to be independent when $P(A \cap B) = P(A)P(B)$.

Part a Therefore, when we say that E is independent of itself,

$$\begin{aligned} P(E \cap E) &= P(E)P(E) \\ P(E) &= P(E)^2 \\ P(E) - P(E)^2 &= 0 \\ P(E)(1 - P(E)) &= 0 \end{aligned}$$

Thus, P(E) = 0 or P(E) = 1.

Part b Using the independence property of A and B,

$$P(A \cap B) = P(A)P(B) = 0.3 * 0.4 = 0.12$$

Therefore, P(A ∪ B) when A and B are independent is

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.3 + 0.4 - 0.12 = 0.58 \end{aligned}$$

We know that when A and B are mutually exclusive sets P(A ∩ B) = 0.
Thus, P(A ∪ B) when A and B are mutually exclusive is

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.3 + 0.4 = 0.7 \end{aligned}$$

Part c Given P(A) = 0.6 and P(B) = 0.8.
Suppose that A and B are independent. Thus, $P(A \cap B) = P(A)P(B) = 0.8 * 0.6 = 0.48$. So $P(A)$ will be as follows,

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.6 + 0.8 - 0.48 = 0.92 \end{aligned}$$

Since, this case follows all the properties of probability, this case is possible.

Now, suppose that A and B are mutually exclusive. Therefore P(A∩B) will be 0 and P(A∪B) will be,

$$\begin{aligned} (A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.6 + 0.8 = 1.4 > 1 \end{aligned}$$

Since P(A∪B) > 1, it breaks a property of probability and thus this case cannot be possible.

Exercise 6
Which of the following are valid CDF's? For each that is not valid, state at least one reason why. For each that is valid, find P($X^2 > 5$).

1. $$F(x) = \begin{cases} e^{-x^2}/4 & if \quad x < 0 \\ 1 - e^{-x^2}/4 & if \quad x \geq 0 \end{cases}$$

2. $$F(x) = \begin{cases} 0 & if \quad x < 0 \\ 0.5 + e^{-x} & if \quad 0 \leq x < 3 \\ 1 & if \quad x \geq 3 \end{cases}$$

3. $$F(x) = \begin{cases} 0 & if \quad x < 0 \\ 0.5 + x/20 & if \quad 0 \leq x \leq 10 \\ 1 & if \quad x \geq 10 \end{cases}$$

Solution: The properties of a valid CDF are that it must be non-decreasing, right continous $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$.

Part 1 The given F satisfies all the conditions to be a valid CDF.

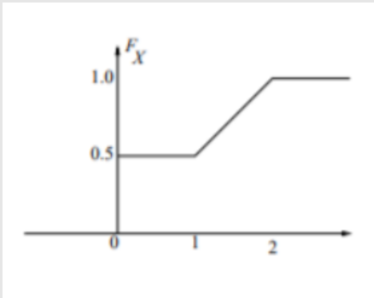
$$\begin{aligned} P(X^2 > 5) &= P(X < -\sqrt{5} \cup X > \sqrt{5}) \\ &= P(X < -\sqrt{5}) + P(X > \sqrt{5}) \\ &= P(X < -\sqrt{5}) + (1 - P(X \leq \sqrt{5})) \\ &= F(-\sqrt{5}) + (1 - F(\sqrt{5})) \\ &= e^{-(-\sqrt{5})^2}/4 + (1 - (1 - e^{-(\sqrt{5})^2}/4)) \\ &= e^{-5}/4 + e^{-5}/4 = e^{-5}/2 \end{aligned}$$

Part 2 This F is not a valid CDF as $0.5 + e^{-x}$ is a decreasing function as its derivative is negative($-e^{-x}$).

Part 3 The given F is a valid CDF as it satisfies all the conditions to be a CDF.

$$\begin{aligned} P(X^2 > 5) &= F(-\sqrt{5}) + (1 - F(\sqrt{5})) \\ &= 0 + (1 - (0.5 + (\sqrt{5})/20)) \\ &= 0.5 - \sqrt{5}/20 \end{aligned}$$

Exercise 7
Let X have the CDF shown.



1. Find P(X $\leq$ 0.8).
2. Find E(X).
3. Find Var(x).

Solution: Since the F_X is a CDF we can define the function as,

$$F_X = \begin{cases} 0, & if \quad x < 0 \\ 0.5, & if \quad 0 \leq x < 1 \\ 0.5x, & if \quad 1 \leq x < 2 \\ 1, & if \quad x \geq 2 \end{cases}$$

Part 1 Therefore, $P(X \leq 0.8) = F_X(0.8) = 0.5$.

Part 2 Since the CDF has a disjoint point but is continous elsewhere shows that it is a mixture of a discrete and a continous random variable.
So, at $x = 0$ the disjoint point, the length of the jump would be the value of the discrete probability at that point,

$$\begin{aligned} P_X(X = 0) &= \lim_{h \to 0^-} (F_X(0) - F_X(h)) \\ &= 0.5 \end{aligned}$$

Using this the PDF can be defined as,

$$f_X = \begin{cases} 0.5, & if \ 1 \leq x < 2 \\ 0, & otherwise \end{cases}$$

Since we know that for discrete random variable $E(X) = \sum_{i=1}^{\infty} x_i P_X(x_i)$ and for continous $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$. So for this CDF we can write,

$$\begin{aligned}
E(X) &= P_X(X = 0) * 0 + \int_1^2 0.5t dt \\
&= 0 + 0.25[t^2]_1^2 \\
&= 0.75
\end{aligned}$$

Part 3 Since $Var(X)$ is defined as $E(X^2) - (E(X))^2$,

$$\begin{aligned}
E(X^2) &= P_X(X = 0) * 0^2 + \int_1^2 0.5t^2 dt \\
&= 0 + \frac{1}{6} * [t^3]_1^2 = \frac{7}{6}
\end{aligned}$$

Thus, Var(X) can be found as

$$\begin{aligned}
Var(X) &= E(X^2) - (E(X))^2 \\
&= \frac{7}{6} - (0.75)^2 \\
&= \frac{7}{6} - \frac{9}{16} = \frac{29}{48} \approx 0.604
\end{aligned}$$

Exercise 8
If the density function of X equals

$$f(x) = \begin{cases} ce^{-2x} & if \quad 0 \leq x < \infty \\ 0 & if \quad x < 0 \end{cases}$$

Find c. What is the value of P(X > 2)?

Solution: The total probability should be equal to 1. Therefore,

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} f(x) dx \\
1 &= \int_0^{\infty} ce^{-2x} dx = \frac{c}{2} \cdot [-e^{-2x}]_0^{\infty} \\
1 &= \frac{c}{2} \\
c &= 2
\end{aligned}$$

Now, to find P(X > 2) we can simply,

$$\begin{aligned}
P(X > 2) = \int_2^{\infty} 2e^{-2x} dx &= [-e^{-2x}]_2^{\infty} \\
&= e^{-4}
\end{aligned}$$

Exercise 9
Suppose a coin having probability 0.7 of coming up heads is tossed three times. Let X denote the number of heads that appear in the three tosses. Determine the probability mass function of X.

Solution: P(H) = 0.7 where H is the event that the coin lands a head. So the PMF of X $\in$ {0, 1, 2, 3} can be found by taking all the possible cases,
Case 1 - 0 Heads

$$\begin{aligned} P(X=0) &= {}^3C_0 P(H^c)^3 \\ &= (0.3)^3 = 0.027 \end{aligned}$$

Case 2 - 1 Head

$$\begin{aligned} P(X=1) &= {}^3C_1 P(H^c)^2 P(H) \\ &= 3 * (0.3)^2 * (0.7) = 0.189 \end{aligned}$$

Case 3 - 2 Heads

$$\begin{aligned} P(X=2) &= {}^3C_2 P(H^c) P(H)^2 \\ &= 3 * (0.3) * (0.7)^2 = 0.441 \end{aligned}$$

Case 4 - 3 Heads

$$\begin{aligned} P(X=3) &= {}^3C_3 P(H)^3 \\ &= (0.7)^3 = 0.343 \end{aligned}$$

Finally the PMF can be shown as,

$$F_X(x) = \begin{cases} 0.027, & if\ x = 0 \\ 0.189, & if\ x = 1 \\ 0.441, & if\ x = 2 \\ 0.343, & if\ x = 3 \\ 0, & otherwise \end{cases}$$

Exercise 10
Let X is a random variable with probability density function

$$f_X(x) = \begin{cases} \text{2x} & if\ 0 \le x \le 1, \\ 0 & otherwise. \end{cases}$$

Find P($X \geq 0.4 | X \leq 0.8$).

Solution: Cumulative probability F_X is,

$$\begin{aligned} F_X(x) = \int_{-\infty}^{x} f_X(t)dt &= \int_0^x f_X(t)dt \\ &= \int_0^x 2tdt = x^2 \end{aligned}$$

Now for P(X $\geq$ 0.4|X $\leq$ 0.8),

$$\begin{aligned} P(X \geq 0.4 | X \leq 0.8) &= \frac{P(X \geq 0.4\ and\ X \leq 0.8)}{P(X \leq 0.8)} \\ &= \frac{F_X(0.8) - F_X(0.4)}{F_X(0.8)} \\ &= \frac{(0.8)^2 - (0.4)^2}{(0.8)^2} = \frac{0.64 - 0.16}{0.64} = \frac{3}{4} = 0.75 \end{aligned}$$

Exercise 11
Let X is an exponentially distributed random variable with parameter λ. For any a,b > 0, find P(X > a + b | X > a).

Solution: Since X is an exponentially distributed random variable with parameter λ, the PDF is thus given by,

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & if x \geq 0 \\ 0, & if x < 0 \end{cases}$$

Thus, $P(X > a)$ can be found out to be,

$$\begin{aligned} P(X > a) = \int_a^\infty f_X(x)dx &= \int_a^\infty \lambda e^{-\lambda x} dx \\ &= [-e^{-\lambda x}]_a^\infty = e^{-\lambda a} \end{aligned}$$

Thus $P(X > a + b|X > a)$ is,

$$\begin{aligned} P(X > a + b|X > a) = \frac{P(X > a + b \; and X > a)}{P(X > a)} &= \frac{P)(X > a + b)}{P(X > a)} \\ &= \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} = e^{-\lambda b} \end{aligned}$$

Exercise 12
Suppose five fair coins are tossed. Let E be the event that all coins land heads. Define a random variable I_E

$$I_E = \begin{cases} 1 & if & E \; occurs \\ 0 & if & E^c occurs \end{cases}$$

For what outcomes in the original sample space does I_E equals 1 ? What is $P\{I_E = 1\}$?

Solution: $I_E = 1$ when all 5 coins land head.

$$\begin{aligned} P(I_E = 1) &= P(H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5) \\ &= \prod_{i=1}^{5} P(H_i)\{H_i \; are \; independent\} \\ &= \left(\frac{1}{2}\right)^5 = \frac{1}{32} \approx 0.03 \end{aligned}$$

Exercise 13
Suppose the distribution function of X is given by

$$F(b) = \begin{cases} 0, & b < 0 \\ \frac{1}{2}, & 0 \leq b < 1 \\ 1, & 1 \leq b < \infty \end{cases}$$

What is the probability mass function of X?

Solution: Since there is a discontinuity at $x = 0, 1$, the PMF should have values only at these points.
So the jump at those points will be the probability of that point,

$$P_X(X = 0) = \lim_{h \to 0^-} (F(0) - F(h)) = 0.5$$

$$P_X(X = 1) = \lim_{h \to 1^-} (F(1) - F(h)) = 0.5$$

Therefore, the PMF of X can be written as,

$$P_X(X = x) = \begin{cases} 0.5, & if\ x = 0 \\ 0.5, & if\ x = 1 \\ 0, & otherwise \end{cases}$$

Exercise 14
A ball is drawn from an urn containing three white and three black balls. After the ball is drawn, it is then replaced and another ball is drawn. This goes on indefinitely. What is the probability that of the first four balls drawn, exactly two are white?

Solution: Let the event white ball drawn be denoted by W and black been drawn by B. Therefore, P(W) = P(B) = 0.5.
Now, for drawing 4 balls of which exactly 2 are white,

$$P(E) = P(W)^2 P(B)^2 = (0.5)^4 = 0.0625$$

But the order of the balls can be changed and thus results in ${}^4C_2$ different cases. Therefore the total probability will be,

$$= {}^4C_2 * P(E) = 6 * 0.0625 = 0.375$$

Exercise 15
A coin having probability p of coming up heads is successively flipped until the r^{th} head appears. Argue that X, the number of flips required, will be n, n $\geq$ r, with probability

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r},\ n \geq r$$

Solution: We stop flipping the coin when the r^{th} head appears, which implies that the last coin should give a head as the output. Therefore, we have $(n-1)$ places where there will be $(n-r)$ tails and the rest heads, thus this probability would be,

$$P(X = n) = p^{r-1}(1-p)^{n-r}$$

But since the positions of the heads can be random and each position is equally likely we multiply by ${}^{n-1}C_{r-1}$ to get the probability,

$$P(X = n) = {}^{n-1}C_{r-1} p^{r-1}(1-p)^{n-r}$$

Now, finally we multiply with p for the last head to occur and thus it gives us the final probability as,

$$P(X = n) = {}^{n-1}C_{r-1} p^r (1-p)^{n-r},\ n \geq r$$

Exercise 16
Suppose that the number of typographical errors on a single page of this book has a Poisson distribution with parameter $\lambda = 1$. Calculate the probability that there is at least one error on this page.

Solution: Let X denote the number of errors on this page. Then since it is a Poisson distribution and the parameter $\lambda = 1$.

$$P(X = n) = \frac{e^{-\lambda}\lambda^n}{n!} = \frac{e^{-1}}{n!}$$

Now for there to be atleast one error or $P(X > 0)$,

$$\begin{aligned} P(X > 0) &= 1 - P(X = 0) \\ &= 1 - \frac{e^{-1}}{0!} \\ &= 1 - e^{-1} \end{aligned}$$