Conditional Probability

Question 1. Of all failures of a certain type of computer hard drive, it is determined that in 20% of them only the sector containing the file allocation table is damaged, in 70% of them only nonessential sectors are damaged, and in 10% of the cases both the allocation sector and one or more nonessential sectors are damaged. A failed drive is selected at random and examined.

a. What is the probability that the allocation sector is damaged?
b. What is the probability that a nonessential sector is damaged?
c. If the drive is found to have a damaged allocation sector, what is the probability that some nonessential sectors are damaged as well?
d. If the drive is found to have a damaged nonessential sector, what is the probability that the allocation sector is damaged as well?
e. If the drive is found to have a damaged allocation sector, what is the probability that no nonessential sectors are damaged?
f. If the drive is found to have a damaged nonessential sector, what is the probability that the allocation sector is not damaged?

Solution: Let us define two events:

$A\ = Event\ of\ the\ allocation\ sector\ is\ damaged.$

$N =\ Event\ of\ the\ nonessential\ sectors\ are\ damaged.$

We are given,

$Probability\ that\ only\ the\ allocation\ sector\ is\ damaged: P(A \cap N^c)\ =\ 0.20$

$Probability\ that\ only\ nonessential\ sectors\ are\ damaged: P(A^c \cap N)\ =\ 0.70$

$Probability\ that\ both\ the\ allocation\ sector\ and\ nonessential\ sectors\ are\ damaged: P(A \cap N) =\ 0.10$

a. The probability that the allocation sector is damaged:

$P(A) =\ P(A \cap\ N^c) +\ P(A \cap\ N) =\ 0.20\ +\ 0.10\ =\ 0.30$

Interpretation: There is a 30% chance that the allocation sector is damaged in a failed hard drive.

b. The probability that a nonessential sector is damaged:

$P(N) =\ P(A^c \cap\ N) +\ P(A \cap\ N) =\ 0.70\ +\ 0.10\ =\ 0.80$

Interpretation: There is an 80% chance that a nonessential sector is damaged in a failed hard drive.

c. If the drive is found to have a damaged allocation sector, the probability that some nonessential sectors are damaged as well:

$$P(N|A) = \frac{P(A \cap\ N)}{P(A)}\ = \frac{0.10}{0.30} = \frac{1}{3}$$

Interpretation: Given that the allocation sector is damaged, there is approximately a 33.3% chance that some nonessential sectors are also damaged.

d. If the drive is found to have a damaged nonessential sector, the probability that the allocation sector is damaged as well:

$$P(A|N) = \frac{P(A \cap N)}{P(N)} = \frac{0.10}{0.80} = \frac{1}{8}$$

Interpretation: Given that a nonessential sector is damaged, there is a 12.5% chance that the allocation sector is also damaged.

e. If the drive is found to have a damaged allocation sector, the probability that no nonessential sectors are damaged:

$$P(N^c|A) = \frac{P(A \cap N^c)}{P(A)} = \frac{0.20}{0.30} = \frac{2}{3}$$

Interpretation: Given that the allocation sector is damaged, there is approximately a 66.7% chance that no nonessential sectors are damaged.

f. If the drive is found to have a damaged nonessential sector, the probability that the allocation sector is not damaged:

$$P(A^c|N) = \frac{P(A^c \cap N)}{P(N)} = \frac{0.70}{0.80} = \frac{7}{8}$$

Interpretation: Given that a nonessential sector is damaged, there is an 87.5% chance that the allocation sector is not damaged.

Question 2. During the monsoon season in a coastal city, there is a 60% chance of rain on any given day. When it rains, there is a 50% chance that strong winds will accompany the rain. On days when it does not rain, there is still a 20% chance that the city will experience strong winds due to nearby storms.

a. What is the probability that it will not rain on a random day during the monsoon season?
b. What is the probability that strong winds will not occur when it does not rain?
c. What is the probability that, on a random day of the monsoon season, it will rain and strong wind?
d. What is the probability that it will not rain and there will be no strong winds on a random day?
e. Suppose, there is a 38% chance that there will be strong winds on a random day regardless of whether it rains or not. Determine the probability that it rained that day if strong winds are observed.

Solution: Let,

$R = Event\ of\ rain\ on\ a\ random\ day$

$W = Event\ of\ wind\ on\ a\ random\ day$

Here, $P(R) = 0.6; P(W|R) = 0.5; P(W|R') = 0.2$

a. Probability that it will not rain on a random day during the monsoon season,

$$P(R') = 1 - P(R) = 0.4$$

Comment: 40% chance that it will not rain on a random day during the monsoon season.

b. The probability that strong winds will not occur given that it does not rain,

$$P(W'|R') = 1 - P(W|R') = 1 - 0.2 = 0.8$$

Comment: 80% chance that strong winds will not occur when it does not rain.

c. The probability that, it will rain and strong wind,

$$P(R \cap W) = P(W \mid R) \times P(R) = 0.5 \times 0.6 = 0.3$$

Comment: 30% chance that it will rain and strong wind.

d. The probability that it will not rain and there will be no strong winds on a random day,

$$P(R' \cap W') = P(W'|R') \times P(R')$$

Here,

$$P(R') = 1 - P(R) = 0.4$$

$$P(W'|R') = 1 - P(W|R') = 1 - 0.2 = 0.8$$

$$\therefore P(R' \cap W') = P(W'|R') \times P(R') = 0.4 \times 0.8 = 0.32$$

Comment: 32% chance that it will not rain and not any strong winds.

e. Here, Probability that there will be strong winds on a random day regardless of whether it rains or not, $P(W) = 0.38$. The probability that it rained that day if strong winds are observed,

$$P(R|W) = \frac{P(R \cap W)}{P(W)} = \frac{0.3}{0.38} = 0.79$$

Comment: Approximately, 79% chance that it rained that day if strong winds are observed.

Question 3. At BRAC University, a survey was conducted among the students of the Department of Computer Science and Engineering. The survey revealed the following statistics: 40% of the students reported difficulties in understanding Algorithms. 30% of the students reported difficulties in understanding Data Structures. 20% of the students reported difficulties in both Algorithms and Data Structures. A student is selected at random from the department.

a. If the student reported difficulties in Data Structures, estimate the probability that they also reported difficulties in Algorithms.

b. If the student reported difficulties in Algorithms, estimate the probability that they also reported difficulties in Data Structures

c. Calculate the probability that the student reported difficulties in either Algorithms or Data Structures.

d. Given that a student did not report difficulties in Algorithms, estimate the probability that they reported difficulties in Data Structures.

e. What is the probability that the student did not report difficulties in both Algorithms and Data Structures?

Solution: Given,

$P(A) = 0.40$ (Probability of difficulties in Algorithms)

$P(B) = 0.30$ (Probability of difficulties in Data Structures)

$P(A \cap B) = 0.20$ (Probability of difficulties in both Algorithms and Data Structures)

a. The conditional probability $P(A \mid B)$ is calculated as:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.20}{0.30} = \frac{2}{3} \approx 0.6667$$

b. The conditional probability $P(B \mid A)$ is calculated as:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.20}{0.40} = \frac{1}{2} = 0.50$$

c. The probability $P(A \cup B)$ is calculated as:

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.40 + 0.30 - 0.20 = 0.50$

d. Let A^c denote the complement of A, the event that the student did not report difficulties in Algorithms.

$P(A^c) = 1 - P(A) = 1 - 0.40 = 0.60$

Using the law of total probability:

$P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c)$

Substituting the known values:

$0.30 = (0.50 \times 0.40) + P(B \mid A^c) \times 0.60$

$0.30 = 0.20 + 0.60P(B \mid A^c)$

$0.10 = 0.60P(B \mid A^c)$

$$P(B \mid A^c) = \frac{0.10}{0.60} = \frac{1}{6} \approx 0.1667$$

e. The probability $P(A^c \cap B^c)$ is calculated as:

$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - 0.50 = 0.50$

Question 4. A survey conducted by an e-commerce company wants to know people's buying preference whether online platforms or traditional platforms among both male and female participants. It is found that 35% of total participants prefer online platforms to purchase. Where, 70% of Participants who prefer online platform were female and 60% of participants who didn't prefer online platform were male. A participant randomly selected

a. What is the probability that the participant prefers traditional platforms?
b. What is the probability that the participant is female and prefers online platforms?
c. What is the probability that the participant is male and prefers online platforms?
d. What is the probability that the participant is female and prefers traditional platforms?
e. If the probability of being a woman is 55 %, what is the probability that she will prefer online platforms?

Solution: Given,

The probability of preferring online platforms is: $P(O) = 0.35$

The probability of being women given that participant prefer online platforms is: $P(F|O) = 0.7$

The probability of being men given that participant didn't prefer online platforms is: $P(M|O') = 0.6$

a. If, O' be the event of traditional platform (or, not preferring the online platform), then the probability that the participant prefers traditional platforms,

$P(O') = 1 - P(O) = 0.65$

b. Probability that the participant is female and prefers online platforms,

$P(F \cap O) = P(F|O) \times P(O) = 0.7 \times 0.35 = 0.245$

c. The probability of being men given that participant prefer online platforms is:

$P(M|O) = 1 - P(F|O) = 0.3$

Probability that the participant is male and prefers online platforms,

$P(M \cap O) = P(M \mid O) \times P(O) = 0.3 \times 0.35 = 0.105$

d. The probability of being female given that participant prefer traditional platforms is,

$P(F|O') = 1 - P(M|O') = 0.4$

Probability that the participant is female and prefers traditional platforms,

$P(F \cap O') = P(F \mid O') \times P(O') = 0.4 \times 0.65 = 0.26$

e. Probability that participant will prefer online platforms given that participant is female,

$$P(O|F) = \frac{P(O \cap F)}{P(F)} = \frac{0.245}{0.55} = 0.445$$

Question 5. In a cultural society of BRACU of 100 students, 70 are girls and 30 are boys. Among the girls, 20 are in the science club, and among the boys, 10 are in the science club.

a. What is the probability that a randomly selected student is a girl?

b. What is the probability that a randomly selected student is in the science club?

c. What is the probability that a student is a girl given that the student is in the science club?

d. What is the probability that a student is in the science club given that the student is a boy?

Solution:

a. Probability that the student is a girl:

$$P(G) = \frac{70}{100} = 0.7$$

b. Probability that the student is in the science club:

Total in the science club = 20 (girls) + 10 (boys) = 30.

$$P(SC) = \frac{30}{100} = 0.3$$

c. Probability that a student is a girl given that the student is in the science club:

$$P(G|SC) = \frac{P(G \cap SC)}{P(SC)} = \frac{20/100}{30/100} = \frac{20}{30} = \frac{2}{3} \approx 0.67$$

d. Probability that a student is in the science club given that the student is a boy:

$$P(SC|B) = \frac{P(SC \cap B)}{P(B)} = \frac{10/100}{30/100} = \frac{10}{30} = \frac{1}{3} \approx 0.33$$

Question 6. The probability that a married man watches a certain TV show is 0.4 and that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does, is 0.7.

a. Find the probability that a married couple watches the show.

b. Find the probability that a wife watches the show given that her husband does.

c. Find the probability that at least one person of a married couple will watch the show.

Solution: Let us define two events H and W as follows:

H: Husband watches the show

W: Wife watches the show

It is given that, $P(H) = 0.4, P(W) = 0.5$ *and* $P(M|W) = 0.7$

a. $P(H \cap W) = P(M|W) \times P(W)$

$= 0.7 \times 0.5$

$= 0.35$

b. $P(W|H) = \frac{P(H \cap W)}{P(H)} = \frac{0.35}{0.4} = 0.875$

c. $P(H \cup W) = P(H) + P(W) - P(H \cap W)$

$$= 0.4 + 0.5 - 0.35$$

$$= 0.55$$

Question 7. The Roads and Highways Department (RHD) conducted a survey during Christmas vacation. The probability that a vehicle entering a hillside tunnel has license plates is 0.33 and the probability that it is a campervan is 0.48. The probability that it is a campervan with license plates is 0.09.

a. What is the probability that a campervan entering the tunnel has license plates?

b. What is the probability that a vehicle with license plates entering the tunnel is a camper van?

c. What is the probability that a vehicle entering the tunnel does not have license plates or is not a campervan?

Solution: Let,

$C = the\ vehicle\ is\ a\ campervan$

$L = the\ vehicle\ has\ license\ plates$

Given that, $P(L) = 0.33, P(C) = 0.48, P(L \cap C) = 0.09$

a. The probability that a campervan entering the tunnel has license plates:

$$P(L|C) = \frac{P(L \cap C)}{P(C)} = \frac{0.09}{0.48} = 0.1875$$

b. The probability that a vehicle with license plates entering the tunnel is a camper van:

$$P(C|L) = \frac{P(L \cap C)}{P(L)} = \frac{0.09}{0.33} = 0.27$$

c. The probability that a vehicle entering the tunnel does not have license plates or is not a campervan:

$$P(L' \cup C') = 1 - P(L \cap C) = 1 - 0.09 = 0.91$$

Question 8. In a city library, a survey was conducted among library visitors. The survey provided the following data: 50% of the visitors reported difficulties in finding books in the Fiction section. 40% of the visitors reported difficulties in finding books in the Non-Fiction section. 25% of the visitors reported difficulties in both the Fiction and Non-Fiction sections. 15% of the visitors reported difficulties in neither the Fiction nor the Non-Fiction sections. A visitor is selected at random from the library.

a. If the visitor reported difficulties in finding books in the Non-Fiction section, estimate the probability that they also reported difficulties in finding books in the Fiction section.

b. If the visitor reported difficulties in finding books in the Fiction section, estimate the probability that they also reported difficulties in finding books in the Non-Fiction section.
c. Calculate the probability that the visitor reported difficulties in either the Fiction or the Non-Fiction section.
d. Given that a visitor did not report difficulties in finding books in the Non-Fiction section, estimate the probability that they reported difficulties in finding books in the Fiction section.
e. What is the probability that the visitor did not report difficulties in either the Fiction or the Non-Fiction sections?

Solution: Let's denote the following events:

F: The event that a visitor reported difficulties in finding books in the Fiction section.

N: The event that a visitor reported difficulties in finding books in the Non-Fiction section.

We are given: $P(F) = 0.50$; $P(N) = 0.40$; $P(F \cap N) = 0.25$; $P(F^c \cap N^c) = 0.15$

a. The probability that a visitor who reported difficulties in the Non-Fiction section also reported difficulties in the Fiction section is:

$$P(F|N) = \frac{P(F \cap N)}{P(N)} = \frac{0.25}{0.40} = 0.625$$

b. The probability that a visitor who reported difficulties in the Fiction section also reported difficulties in the Non-Fiction section is:

$$P(N|F) = \frac{P(F \cap N)}{P(F)} = \frac{0.25}{0.50} = 0.50$$

c. To find the probability that a visitor reported difficulties in either the Fiction or the Non-Fiction section:

$$P(F \cup N) = P(F) + P(N) - P(F \cap N)$$
$$P(F \cup N) = 0.50 + 0.40 - 0.25 = 0.65$$

d. To find the probability that a visitor who did not report difficulties in the Fiction section reported difficulties in the Non-Fiction section:

$$P(F^c) = 1 - P(F) = 1 - 0.50 = 0.50$$
$$P(N \cap F^c) = P(N) - P(F \cap N) = 0.40 - 0.25 = 0.15$$
$$P(N|F^c) = \frac{P(N \cap F^c)}{P(F^c)} = \frac{0.15}{0.50} = 0.30$$

e. To find the probability that the visitor did not report difficulties in either the Fiction or the Non-Fiction section:

$$P(F^c \cap N^c) = 0.15$$

So, the probability that the visitor did not report difficulties in either the Fiction or the Non-Fiction section is 0.15.

Bayes Theorem

Question 1. A certain delivery service offers both express and standard delivery. Seventy-five percent of parcels are sent by standard delivery, and 25% are sent by express. Of those sent standard, 80% arrive the next day, and of those sent express, 95% arrive the next day. A record of a parcel delivery is chosen at random from the company's files.

a. What is the probability that the parcel was shipped express and arrived the next day?
b. What is the probability that it arrived the next day?
c. Given that the package arrived the next day, what is the probability that it was sent express?

Solution: Let us define the following events:

$S = \textit{the parcel is sent by standard delivery.}$

$E = \textit{the parcel is sent by express delivery.}$

$A = \textit{the parcel arrives the next day.}$

Given, Probability that a parcel is sent by standard delivery: $P(S) = 0.75$

Probability that a parcel is sent by express delivery: $P(E) = 0.25$

Probability that a parcel sent by standard delivery arrives the next day: $P(A \mid S) = 0.80$

Probability that a parcel sent by express delivery arrives the next day: $P(A \mid E) = 0.95$

a. The probability that the parcel was shipped express and arrived the next day:

   This is the joint probability of occurring both events (E) and (A):

   $P(E \cap A) = P(E) \times P(A \mid E) = 0.25 \times 0.95 = 0.2375$

   Interpretation: There is a 23.75% chance that a parcel was shipped express and arrived the next day.

b. The probability that it arrived the next day: This is the total probability of happening (A), considering both standard and express deliveries:

   $P(A) = P(S) \times P(A \mid S) + P(E) \times P(A \mid E)$

   $= 0.75 \times 0.80 + 0.25 \times 0.95$

   $= 0.8375$

   Interpretation: There is an 83.75% chance that a randomly chosen parcel from the company's files arrived the next day.

c. Given that the package arrived the next day, the probability that it was sent express:

   $$P(E \mid A) = \frac{P(E \cap A)}{P(A)} = \frac{0.2375}{0.8375} = 0.2836$$

   Interpretation: Given that a parcel arrived the next day, there is approximately a 28.36% chance that it was sent by express delivery.

Question 2. A certain level of customer dissatisfaction is present in about 1 out of 1000 customers in a given population. Suppose there is a customer feedback survey, and customers are asked to rate their satisfaction level. The survey results show that a positive rating is given with a probability of 10% for a dissatisfied customer and with a probability of 95% for a satisfied customer.

a. If a customer is selected at random from this population, what is the probability that the feedback survey will result in a positive rating? Interpret your findings.

b. If the feedback survey of a selected customer results in a negative rating, what is the probability that they are actually dissatisfied?

Solution: Let,

$D = Event\ of\ dissatisfied\ customer$

$S = Event\ of\ satisfied\ customer$

$PR = Event\ of\ positive\ rating$

$NR = Event\ of\ negative\ rating$

a. The probability that the feedback survey will result in a positive rating,

$P(PR) = P(PR|D) \times P(D) + P(PR|S) \times P(S)$

Here, $P(D) = \frac{1}{1000} = 0.001$; $P(S) = 1 - P(D) = 0.999$; $P(PR|D) = 0.10$; $P(PR|S) = 0.95$

$\therefore\ P(PR) = P(PR|D) \times P(D) + P(PR|S) \times P(S)$

$= (0.10 \times 0.001) + (0.999 \times 0.95) = 0.95$

b. The probability that they are actually dissatisfied when selected customer results in a negative rating, $P(D|NR) = (P(NR|D) \times P(D))/P(NR)$

$\therefore\ P(NR) = 1 - P(PR) = 1 - 0.95 = 0.05$

$\therefore\ P(NR|D) = 1 - P(PR|D) = 1 - 0.10 = 0.90$

$\therefore\ P(D) = 0.001$

$\therefore\ P(D|NR) = (P(NR|D) \times P(D))/P(NR) = ((0.90 \times 0.001))/0.05 = 0.018$

Question 3. A company has developed a new drug test that can detect a particular substance in a person's blood. The test has the following characteristics: **True Positive Rate (Sensitivity):** 97% (the probability that the test correctly identifies a person with the substance as positive). **False Positive Rate:** 5% (the probability that the test incorrectly identifies a person without the substance as positive). It is known that 1% of the population uses this substance. A person is randomly selected from the population and tests positive. **What is the probability that this person actually uses the substance?**

Solution: Let's define the events:

U: The person uses the substance.

U': The person does not use the substance.

T^+: The test result is positive.

We want to find the probability that a person actually uses the substance given a positive test result, $P(U \mid T^+)$. By **Bayes' Theorem**:

$$P(U \mid T^+) = \frac{P(T^+ \mid U) \cdot P(U)}{P(T^+)}$$

Step 1: Find Each Component for Bayes' Theorem

- **Prior Probability:** The probability that a randomly selected person uses the substance:
$$P(U) = 0.01$$
- **Complement Probability:** The probability that a randomly selected person does not use the substance:
$$P(U') = 1 - P(U) = 0.99$$
- **True Positive Rate (Sensitivity):** The probability that the test is positive given the person uses the substance:
$$P(T^+ \mid U) = 0.97$$
- **False Positive Rate:** The probability that the test is positive given the person does not use the substance:
$$P(T^+ \mid U') = 0.05$$
- **Total Probability of Testing Positive (Law of Total Probability):**
$$P(T^+) = P(T^+ \mid U) \cdot P(U) + P(T^+ \mid U') \cdot P(U')$$

Substitute the values:

$$P(T^+) = (0.97 \times 0.01) + (0.05 \times 0.99)$$
$$P(T^+) = 0.0097 + 0.0495 = 0.0592$$

Step 2: Apply Bayes' Theorem

Now substitute these values into Bayes' Theorem:

$$P(U \mid T^{+}) = \frac{P(T^{+} \mid U) \cdot P(U)}{P(T^{+})}$$

$$P(U \mid T^{+}) = \frac{0.97 \times 0.01}{0.0592}$$

$$P(U \mid T^{+}) = \frac{0.0097}{0.0592} \approx 0.1638$$

Final Answer: The probability that a person who tests positive actually uses the substance is approximately **16.38%**.

Interpretation: Even though the test has a high sensitivity (97%) and a relatively low false positive rate (5%), the probability that a person who tests positive is actually using the substance is still quite low (16.38%). This is due to the low prevalence (1%) of substance use in the population, demonstrating the significance of considering base rates in diagnostic testing scenarios.

Question 4. During the recent student protests in Bangladesh, a survey was conducted among the students of BRAC University. The survey revealed that: 25% of the students physically participated in the protests. Among those who physically participated in the protests, 60% were active in organizing online campaigns. Among those who did not physically participate in the protests, 80% were also active in organizing online campaigns.

a. What is the prior probability that a randomly selected student physically participated in the protests?
b. Given that a student physically participated in the protests, what is the probability that they were active in organizing online campaigns?
c. If a student was active in organizing online campaigns, what is the probability that they actually physically participated in the protests?

Solution: Let,

$P(P)$ be the probability that a student physically participated in the protests.

$P(O \mid P)$ be the probability that a student was active in organizing online campaigns given they participated in the protests.

$P(O \mid P^{c})$ be the probability that a student was active in organizing online campaigns given they did not participate in the protests.

$P(O)$ be the probability that a student was active in organizing online campaigns.

$P(P \mid O)$ be the probability that a student physically participated in the protests given they were active in organizing online campaigns.

a. The prior probability that a randomly selected student physically participated in the protests is given as:

$$P(P) = 0.25$$

b. The probability that a student who physically participated in the protests was active in organizing online campaigns is:

$$P(O \mid P) = 0.60$$

c. To find the probability that a student physically participated in the protests given they were active in organizing online campaigns, we use Bayes' theorem:

$$P(P \mid O) = \frac{P(O \mid P) \cdot P(P)}{P(O)}$$

First, calculate $P(O)$ using the law of total probability:

$$P(O) = P(O \mid P) \cdot P(P) + P(O \mid P^c) \cdot P(P^c)$$

$$P(P^c) = 1 - P(P) = 1 - 0.25 = 0.75$$

$$P(O) = (0.60 \cdot 0.25) + (0.80 \cdot 0.75)$$

$$P(O) = 0.15 + 0.60 = 0.75$$

Now, calculate $P(P \mid O)$:

$$P(P \mid O) = \frac{P(O \mid P) \cdot P(P)}{P(O)}$$

$$P(P \mid O) = \frac{0.60 \cdot 0.25}{0.75} = \frac{0.15}{0.75} = 0.20$$

Thus, the probability that a student physically participated in the protests given they were active in organizing online campaigns is 0.20 or 20%.

Question 5. Two factories supply light bulbs to the market. Bulbs from factory X work for over 5000 hours in 99% of cases, whereas bulbs from factory Y work for over 5000 hours in 95% of cases. It is known that factory X supplies 60% of the total bulbs available in the market.

a. What is the probability that a purchased bulb will work for longer than 5000 hours?

b. Given that a light bulb works for more than 5000 hours, what is the probability that it was supplied by factory Y?

c. Given that a light bulb work does not work for more than 5000 hours, what is the probability that it was supplied by factory X?

Solution: Let, H = bulb works over 5000 hours; X = A bulb comes from factory X; Y = A bulb comes from factory Y

a. $P\,(H) = P(H|X)P(X) + P(H|Y)P(Y) = (0.99)(0.6) + (0.95)(0.4) = 0.974$

b. $P\,(Y|H) = \frac{P(H|Y)\,P(Y)}{P(H)} = \frac{0.95 \times 0.4}{0.974} = 0.39$

c. $P(X|HC) = \frac{P(H^c|X)\,P(X)}{P(H^c)} = \frac{P(H^c|X)\,P(X)}{1-P(H)} = \frac{(1-0.99)\,(0.6)}{1-0.974} = 0.23$

Question 6. A research team working on finding a treatment for a newly transmitted virus. Where, 60% of patients were given medicine and a placebo was given to others. From the treatment group 70% and from the placebo group 30% patient recovered within 1 month.

a. What is the probability that placebo was given and the patient will be recovered?
b. What is the probability that a patient recovered from it?
c. Given that a patient recovered, what is the probability that patient belong to medicine group?

Solution: Given that,

The probability that medicine given to a patient, P(M) = 0.6

The probability that placebo given to a patient, P(P) = 0.4

The probability that medicine given to a patient and the patient recovered, P(R|M) = 0.7

The probability that placebo given to a patient and the patient recovered, P(R|P) = 0.3

a. The probability that placebo was given and the patient will be recovered,

$$P(R \cap P) = P(R|P) \times P(P) = 0.3 \times 0.4 = 0.12$$

b. The probability that a patient recovered from it,

$$P(R) = P(R|M) \times P(M) + P(R|P) \times P(P) = 0.7 \times 0.6 + 0.3 \times 0.4 = 0.54$$

c. Given that patient recovered, the probability that patient belong to medicine group,

$$P(M|R) = \frac{P(M \cap R)}{P(R)} = P(R|M) \times \frac{P(M)}{P(R)} = \frac{0.7 \times 0.6}{0.54} = 0.78$$

Question 7. A blood test gives 96% positive (+) results for diseased people and 3% positive results for healthy people. In a community 2% people have diseases. A person is randomly selected from the community and gives the test.

a. What is the probability that the test will be positive?
b. If the test is positive, what is the probability that the person has the disease?
c. If the test is positive, what is the probability that the person has not the disease?

Solution: Let us define, D for disease; H for healthy; + for positive test result.

It is given that, $P(+|D) = 0.96, P(+|H) = 0.03, P(D) = 0.02, P(H) = 0.98$

a. $P(+) = P(D \cap +) + P(H \cap +) = P(D).P(+|D) + P(H).P(+|H)$

$= 0.02 \times 0.96 + 0.98 \times 0.03$

$= 0.0486$

b. $P(D|+) = \frac{P(D\cap+)}{P(+)} = \frac{P(D)\times P(+|D)}{P(+)} = \frac{0.02\times0.96}{0.0486} = 0.3951$

c. $P(H|+) = 1 - P(D|+) = 1 - 0.3951 = 0.6049$

Question 8. A lie-detector machine has the property that 90% of the guilty suspects are properly judged while 10% of the guilty suspects are improperly found innocent. On the other hand, innocent suspects are misjudged 1% of the time. Suppose, only 5% suspects have ever committed a crime.

a. What is the probability that the suspect is innocent?

b. What is the probability that the suspect is judged guilty?

c. If the machine indicates that a selected suspect is guilty, what is the probability that he is innocent?

Solution: Here the defined events, G=Suspect is guilty**;** I=Suspect is innocent**;** JG= Suspect is judged guilty**;** JI=Suspect is judged innocent

Given that, $P(JG|G) = 0.90, P(JI|G) = 0.10, P(JG|I) = 0.01, P(G) = 0.05$

a. $P(I) = 1 - P(G) = 1 - 0.05 = 0.95$

b. $P(JG) = P(JG|I)P(I) + P(JG|G)P(G) = 0.01 \times 0.95 + 0.90 \times 0.05 = 0.0545$

c. $P(I|JG) = \frac{P(JG|I)P(I)}{P(JG|I)P(I)+P(JG|G)P(G)} = \frac{0.01\times0.95}{0.01\times0.95+0.90\times0.05} = 0.1743$

Question 9.

In a bustling city, there are two popular types of local cafes: artisanal coffee shops and casual diners. The following statistics are available from a recent citywide survey: 70% of the cafes are artisanal coffee shops, while 30% are casual diners. 10% of the artisanal coffee shops offer live music performances. 25% of the casual diners offer live music performances. A city guidebook randomly selects a cafe known for its live music performances.

a. What is the probability that a randomly selected cafe with live music is an artisanal coffee shop?

b. Given that a cafe is known to be an artisanal coffee shop, what is the probability that it offers live music performances?

c. What is the probability that a randomly selected cafe is a casual diner given that it offers live music performances?

Solution:

Let's denote the following events:

A: The event that a cafe is an artisanal coffee shop.

C: The event that a cafe is a casual diner.

M: The event that a cafe offers live music performances.

We are given: $P(A) = 0.70$; $P(C) = 0.30$; $P(M \mid A) = 0.10$ (probability of offering live music given that the cafe is an artisanal coffee shop); $P(M \mid C) = 0.25$ (probability of offering live music given that the cafe is a casual diner)

a. Probability that a Cafe with Live Music is an Artisanal Coffee Shop:

Using Bayes' Theorem:

$$P(A|M) = \frac{P(M|A) \cdot P(A)}{P(M)}$$

First, we need to find P(M), the total probability of a cafe offering live music:

$$P(M) = P(A) \cdot P(M|A) + P(C) \cdot P(M|C) = 0.145$$

compute P(A|M):

$$P(A|M) = \frac{0.10 \times 0.70}{0.145} \approx 0.483$$

So, the probability that a cafe with live music is an artisanal coffee shop is approximately 0.483.

b. Probability that an Artisanal Coffee Shop Offers Live Music:

This is directly given by:

$$P(M \mid A) = 0.10$$

So, the probability that an artisanal coffee shop offers live music is 0.10.

c. Probability that a Cafe is a Casual Diner Given that it Offers Live Music:

Using Bayes' Theorem:

$$P(C|M) = \frac{P(M|C) \cdot P(C)}{P(M)} = 0.517$$

So, the probability that a cafe with live music is a casual diner is approximately 0.517.

<u>**Random Variables**</u>

Question 1. You are participating in a game where you have two chances to throw a ball toward a target hole. The rules of the game are as follows: If no balls land in the hole, you lose 10 Taka. If exactly one ball lands in the hole, you neither gain nor lose anything. If both balls land in the hole, you win 50 Taka. It is known that only 10% of the balls land in the holes, and the throws are independent.

a. Construct a probability mass function (PMF) considering the gain or loss in the game as a random variable.
b. Calculate the expected gain or loss, and determine if playing the game is favorable.
c. Find the standard deviation of the gain or loss.

Solution:

a. Let be X the random variable representing the gain or loss in Taka. We have three possible outcomes:

No balls land in the hole: Loss 10 Taka; and the probability:

$P(X=-10) = 2C0 \times (0.1)^0 \times (1-0.1)^2 = 0.81$ (Using binomial distribution)

Exactly one ball lands in the hole: No gain or loss (0 Taka); and the probability:

$P(X=0) = 2C1 \times (0.1)^1 \times (1-0.1)^1 = 0.18$

Both balls land in the hole: Win 50 Taka; and the probability:

$P(X=50) = 2C2 \times (0.1)^2 \times (1-0.1)^0 = 0.01$

Therefore, the probability mass function (PMF) is:

$\boldsymbol{X=x}$	-10	0	50
$\boldsymbol{p(x)}$	0.81	0.18	0.01

b. Expected Gain or Loss:

The expected value ($E(X)$) is calculated as:

$E(X) = \Sigma x.p(x)$

$= (-10) \times 0.81 + 0 \times 0.18 + 50 \times 0.01$

$= -8.1 + 0 + 0.5 = -7.6$

So, the expected loss is 7.6 Taka. Since the expected value is negative, playing the game is not favorable.

c. Standard Deviation:

First, calculate the variance ($Var\ (X)$):

$Var(X) = E(X^2) - [E(X)]^2$

Where:

$E(X^2) = \Sigma x^2 . p(x)$

$= (-10)^2 \times 0.81 + 0^2 \times 0.18 + 50^2 \times 0.01$

$= 81 + 0 + 25 = 106$

$\therefore Var(X) = 106 - (-7.6)^2$

$= 106 - 57.76 = 48.24$

Finally, the standard deviation $(\sigma(X))$ is:

$$\sigma(X) = \sqrt{48.24} \approx 6.95$$

Thus, the standard deviation of the gain or loss is approximately 6.95 Taka.

Question 2. Suppose that 2 batteries are randomly chosen from a box containing 10 batteries of which 7 are good and 3 are defective. Let X denote the number of defective batteries chose.

a. Summarize the probability function of X.

b. Calculate the probability of

  i. Exactly one defective battery chose

  ii. More than or equal to one defective battery chose

c. Determine the mean and standard deviation of this probability function.

Solution:

a. Here, $\boldsymbol{X}$ denote the number of defective batteries chose which values are 0, 1, and 2.

Probability of 0 defective battery, 2 good batteries: $P(0) = \frac{3C0 \times 7C2}{10C2} = \frac{24}{45}$

Probability of 1 defective battery, 1 good battery: $P(1) = \frac{3C1 \times 7C1}{10C2} = \frac{42}{45}$

Probability of 2 defective battery, 0 good battery: $P(2) = \frac{3C2 \times 7C0}{10C2} = \frac{21}{45}$

Thus, the probability function of X

$\boldsymbol{X}$	0	1	2
$\boldsymbol{P(x)}$	$\frac{24}{45}$	$\frac{42}{45}$	$\frac{21}{45}$

b. $P(Exactly\ one\ defective\ battery\ chose) = P(1) = \frac{42}{45}$

$P(X \geq 1) = P(1) + P(2) = ???$

c. Mean, $E(X) = (0 \times \frac{24}{45}) + (1 \times \frac{42}{45}) + (2 \times \frac{21}{45}) = ???$

Comment:

Variance, $V(X) = E(X^2) - \{E(X)\}^2 = ???$

Comment:

Question 3. A random variable X has the following probability distribution:

$$P(X = k) = \begin{cases} \frac{k}{10}; for\ k = 1,2,3,4 \\ 0;\ Otherwise \end{cases}$$

a. Verify if this is a valid probability distribution.

b. Find $P(X > 1)$ and $P(X < 4)$

c. Calculate the expected value E(X). Calculate E(X²).

d. Calculate the variance Var(X)

Solution:

a. Verify if this is a valid probability distribution:

$$P(1) + P(2) + P(3) + P(4) = \frac{1}{10} + \frac{2}{10} + \frac{3}{10} + \frac{4}{10} = 1\ (valid)$$

b.

$$P\ (X > 1)\ =\ 1 - P(X = 0) - P(X = 1)\ =\ 1 - \frac{0}{10} - \frac{1}{10} = \frac{9}{10}$$

$$P\ (X < 4)\ =\ 1 - P(X = 4)\ =\ 1 - \frac{4}{10} = \frac{6}{10}$$

c. Calculate the expected value E(X):

$$E(X) = \sum kP(X = k) = 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{4}{10} = \frac{1 + 4 + 9 + 16}{10} = \frac{30}{10} = 3$$

Calculate E(X²):

$$E(X^2) = 1^2 \cdot \frac{1}{10} + 2^2 \cdot \frac{2}{10} + 3^2 \cdot \frac{3}{10} + 4^2 \cdot \frac{4}{10} = \frac{1 + 8 + 27 + 64}{10} = \frac{100}{10} = 10$$

d. Calculate the variance Var(X):

$$Var(X) = E(X^2) - \left(E(X)\right)^2 = 10 - 3^2 = 10 - 9 = 1$$

Question 4. The probability function of a discrete random variable X is

$$P(X = x) = \begin{cases} \alpha\left(\frac{3}{4}\right)^x ; x = 0,1,2,3 \\ 0; Elsewhere \end{cases}$$

a. What is the value of α ?

b. Find the value of $P\ (X\ \leq\ 3)$.

c. Calculate the value of $E(X)$ and $E(X^2)$.

Solution:

a. Since f (x) is a probability function, $\sum_x P(x) = 1$

But, $P(0) = \alpha\left(\frac{3}{4}\right)^0 = \alpha,\quad P(1) = \alpha\left(\frac{3}{4}\right)^1 = \alpha\left(\frac{3}{4}\right),\quad P(2) = \alpha\left(\frac{3}{4}\right)^2 = \alpha\left(\frac{9}{16}\right)$; and $P(3) =$

$\alpha\left(\frac{3}{4}\right)^3$ and so on. Hence,

$$1 = \alpha + \alpha\left(\frac{3}{4}\right) + \alpha\left(\frac{3}{4}\right)^2 + \alpha\left(\frac{3}{4}\right)^3$$

$$\Rightarrow 1 = \alpha\left(1 + \frac{3}{4} + \left(\frac{3}{4}\right)^2 + \left(\frac{3}{4}\right)^3\right)$$

$$\Rightarrow 1 = 2.73\alpha$$

$$\Rightarrow \alpha = 0.37$$

b. Hence the complete *pmf* of X is $P(x) = 0.37\left(\frac{3}{4}\right)^x,\quad x = 0, 1, 2, 3$

Also, $P\ (X\ \leq\ 2) = 0.37 + 0.37\left(\frac{3}{4}\right) + 0.37\left(\frac{9}{16}\right) = 0.84$

$$E(X) = \sum xP(X = x) = 0(P(X = 0)) + 1\,(P(X = 1)) + 2(P(X = 2)) + 3\,(P(X = 3))$$

$$= 0 + \frac{3}{4} + 2\left(\frac{3}{4}\right)^2 + 3\left(\frac{3}{4}\right)^3 = ???$$

$E(X^2) = \Sigma x^2 P(X = x) = ???$

Question 5. At BRAC University, students enrolled in the STA 201 course are required to submit 4 assignments throughout the semester. Let X be the discrete random variable representing the number of assignments submitted on time by a student out of the 4 total assignments. The probability mass function (PMF) of X is given by:

$$P(X=x) = \begin{cases} 0.05; & if\ x=0 \\ 0.20; & if\ x=1 \\ 0.35; & if\ x=2 \\ 0.35; & if\ x=3 \\ 0.05; & if\ x=4 \end{cases}$$

If a student is randomly selected, then:

a. Determine the probability that exactly 3 assignments are submitted on time.

b. Compute the expectation (mean) of the random variable X. Based on the computed mean, can you determine whether a randomly selected student is likely to get full marks in assignments, given that full marks are awarded if at least 3, (n - 1), assignments are submitted on time?

c. Compute the variance of the random variable X

d. Suppose your instructor counts an additional assignment for every assignment submitted (e.g., if you submit 2 assignments, it will be counted as 3). Define a new random variable Y to represent this adjusted number of assignments. Determine the expectation and variance of Y. Based on the new mean, can you determine whether a randomly selected student is likely to get full marks in assignments? Also, find the probability mass function (PMF) of Y.

Solution:

a. To determine the probability that exactly 3 assignments are submitted on time, we use the PMF provided:

$$P(X=3) = 0.35$$

b. To compute the expectation (mean) of the random variable X, we use the formula for the expected value of a discrete random variable:

$$E(X) = \sum_{x=0}^{4} x\,P(X=x)$$

Substituting the given PMF values:

$E(X) = (0 \cdot 0.05) + (1 \cdot 0.20) + (2 \cdot 0.35) + (3 \cdot 0.35) + (4 \cdot 0.05)$

$E(X) = 0 + 0.20 + 0.70 + 1.05 + 0.20$

$E(X) = 2.15$

A student is likely to get full marks if they submit at least 3 assignments on time. Given that $E(X) = 2.15$, which is less than 3, it is not very likely that a randomly selected student will get full marks.

c. To compute the variance of X, we first need $E(X^2)$:

$$E(X^2) = \sum_{x=0}^{4} x^2\, P(X = x)$$

$$E(X^2) = (0^2 \cdot 0.05) + (1^2 \cdot 0.20) + (2^2 \cdot 0.35) + (3^2 \cdot 0.35) + (4^2 \cdot 0.05)$$

$$E(X^2) = 0 + 0.20 + 4 \cdot 0.35 + 9 \cdot 0.35 + 16 \cdot 0.05$$

$$E(X^2) = 0.20 + 1.40 + 3.15 + 0.80$$

$$E(X^2) = 5.55$$

The variance is given by:

$$Var(X) = E(X^2) - \left(E(X)\right)^2$$

$$Var(X) = 5.55 - (2.15)^2$$

$$Var(X) = 5.55 - 4.6225$$

$$Var(X) = 0.9275$$

d. Suppose the instructor counts an additional assignment for every assignment submitted, i.e., if a student submits X assignments, it is counted as $Y = X + 1$. Then:

$$E(Y) = E(X + 1) = E(X) + 1$$

$$E(Y) = 2.15 + 1 = 3.15$$

With the new mean $E(Y) = 3.15$, a randomly selected student is likely to get full marks in assignments if the new mean number of assignments (including the additional count) is at least 3. Since $E(Y)$ is 3.15, which is greater than 3, on average, a randomly selected student will get full marks under the new counting scheme.

The variance of Y is the same as the variance of X because adding a constant does not affect the variance:

$$Var(Y) = Var(X) = 0.9275$$

The PMF of Y can be derived from the PMF of X by shifting the values by 1:

$$P(Y = y) = P(X = y - 1)$$

Thus, the PMF of Y is:

$$P(Y = y) = \begin{cases} 0.05; & if\ y = 1 \\ 0.20; & if\ y = 2 \\ 0.35; & if\ y = 3 \\ 0.35; & if\ y = 4 \\ 0.05; & if\ y = 5 \end{cases}$$

Question 6. Suppose in a game, a participant gets two chances to draw a token each time from a gamble box. The participant receives prizes according to the token they draw. There are 5 tokens for special prizes and 8 tokens for ordinary prizes. Let $\boldsymbol{X}$ denote the number of tokens for special prizes chose.

a. Summarize the probability function of $\boldsymbol{X}$.
b. Calculate the probability of exactly two special tokens chose
c. Determine the mean and standard deviation of this probability function.

Solution:

a. Here, $\boldsymbol{X}$ denote the number of tokens for special prizes chose which values are 0, 1, and 2.

Probability of 0 special token, 2 ordinary token: $P(0) = \frac{5C0 \times 8C2}{13C2} = \frac{28}{78}$

Probability of 1 special token, 1 ordinary token: $P(1) = \frac{5C1 \times 8C1}{13C2} = \frac{40}{78}$

Probability of 2 special token, 0 ordinary token: $P(2) = \frac{5C2 \times 8C0}{13C2} = \frac{10}{78}$

Thus, the probability function of X

X	0	1	2
P(X)	$\frac{28}{78}$	$\frac{40}{78}$	$\frac{10}{78}$

b. Exactly one special token chose,

$$P\,(Exactly\ one\ special\ token\ chose) = P\,(1) = \frac{40}{78}$$

c. Mean, $E(X)$ =???

Standard deviation,

<u>Hints:</u>

$V(X) = E(X^2) - (E(X))^2$

$E(X^2) = 0^2.\frac{28}{78} + 1^2.\frac{40}{78} + 2^2.\frac{10}{78}$

Question 7. Let X be a random variable with probability distribution

x	0	1	2	3
$p(x)$	1/3	1/2	0	1/6

Determine,

a. $P(X > 1)$ [Ans: 1/6]

b. $P(X \leq 2)$ [Ans: 5/6]

c. $E(X)$ [Ans: 1]

d. $E(X^2)$ [Ans: 2]

e. $E[(X-1)^2]$ [Ans: 1]

f. $Var(X)$ [Ans: 1]

Question 8. Katrina Kaif's Kay Beauty brand estimates the net profit on a new skincare product. This versatile brand would like to launch this product and try to estimate its profit during the 1st year. If its net profit is \$5 million, it is `popular'; if it is \$2 million, it is 'fairly popular' and a loss of \$2 million if it is `flop'. The brand assigns the following probability to first year expectations for the product; popular: 0.18, fairly popular: 0.45, and flop: 0.37. What are the expectations and standard deviation of the first-year net profit for this new product?

a. Construct a probability mass function (PMF) considering the 1st year net profit of the required brand. Is it a valid probability distribution?

b. Estimate the probability of net profit less than or equal to \$2 million.

c. Determine the mean and standard deviation of net profit for this new product.

d. Calculate E(2X+3) and Var(4X+7).

Solution: Let, X be the random variable that represents net profit during 1st year.

a. Probability mass function (PMF) based on the above situation:

X=x	-2	2	5
P(x)	0.37	0.45	0.18

This PMF will be valid if it satisfies $\sum P(x_i) = 1$

Here, $P(X=-2) + P(X=2) + P(X=5) = 0.37 + 0.45 + 0.18 = 1$

Therefore, the above PMF is valid.

b. $P(X \leq 2) = P(X=-2) + P(X=2) = 0.37 + 0.45 = 0.82$

c. $E(X) = \sum xP(X=x) = (-2)P(X=-2) + (2)P(X=2) + (5)P(X=5)$

$= (-2 \times 0.37) + (2 \times 0.45) + (5 \times 0.18) = 1.06$

$E(X^2) = \sum x^2\, P(X=x) \;= (-2)^2 P(X=-2) + (2)^2 P(X=2) + (5)^2 P(X=5)$

$= (4 \times 0.37) + (4 \times 0.45) + (25 \times 0.18) = 7.78$

$$Var(X) = E(X^2) - \left(E(X)\right)^2 = 7.78 - (1.06)^2 = 6.6564$$

$$SD(X) = \sqrt{Var(X)} = \sqrt{6.66} = 2.58$$

d. $E(2X + 3) = 2E(X) + 3 = 2 \times 1.06 + 3 = 5.12$

$$Var(4X + 7) = 16Var(X) = 16 \times 6.6564 = 106.5024$$

Question 9. Bangladesh Bank has six tellers available to serve customers. The number of tellers busy with customers at peak time, say, 11:00 a.m. varies from day to day. So, it is a random variable denoted by X. It is known from the past records that the probability distribution of X is as follows.

Value of X	0	1	2	3	4	5	6
P(x)	0.03	0.05	0.08	0.15	K	0.26	0.22

a. Find the value of K.
b. Find the expected number of tellers busy with the customers at 11.00 a.m.
c. Find the variance and coefficient of variation of the number of tellers busy with the customers at 11.00 a.m.
d. Suppose the number of tellers busy with customers doubled and the resulting random variable is donated by $Y = 2X$. Determine the variance and CV of Y and express it as a function of X.

Solution:

a. To find K, we use the fact that the sum of all probabilities must equal 1:

$$0.03 + 0.05 + 0.08 + 0.15 + K + 0.26 + 0.22 = 1$$

$$K = 0.21$$

b. The expected value E(X) is calculated as follows:

$$E(x) = \sum_{x} x \cdot P(x) = 3.72$$

c. The variance is calculated as

$$V(x) = E(X^2) - \left(E(X)\right)^2 = 6.69$$

d. The coefficient of variation is

$$CV = \frac{\sqrt{V(x)}}{E(x)} = 0.696$$

e. Try yourself, the answer is: 26.79.

Binomial Distribution

Question 1. Suppose that the probability that an adult person is suffering from diabetes in a population is 0.10. Suppose that we randomly select a sample of 6 persons.

a. Find the probability distribution of the random variable (X) representing the number of persons with diabetes in the sample.
b. Find the expected number of persons with diabetes in the sample (mean of X). [Ans: 0.6]
c. Find the variance of X. [Ans: 0.54]
d. What is the probability that there will be exactly 2 persons with diabetes? [Ans: 0.09842]
e. What is the probability that there will be at most 2 persons with diabetes? [Ans: 0.98415]
f. What is the probability that there will be at least 4 persons with diabetes? [Ans: 0.00128]

Solution: We are interested in the following random variable:

X = the number of persons with diabetes in the sample of 6 persons.

Notes:

i. Bernoulli trial: diagnosing whether a person has diabetes or not. There are two outcomes for each trial
ii. Number of trials = 6, or n=6.
iii. Probability of success: = 0.10.
iv. Probability of failure:
v. The trials are independent because of the fact that the result of each person does not affect the result of any other person since the selection was made at random.

The random variable, X, has a binomial distribution with parameters: n=6 and p=0.10, that is: *X* ~ Binomial (n, p),

X ~ Binomial (6, 0.10).

The possible values of X are X=x, where, x = 0, 1, 3, 4, 5, 6

[Try yourself]

Question 2. Heart failure is due to either natural occurrences (87%) or outside factors (13%). Outside factors are related to induced substances or foreign objects. Natural occurrences are caused by arterial blockage, disease, and infection. Suppose that 20 patients will visit an emergency room with heart failure. Assume that causes of heart failure between individuals are independent.

a. What is the probability that three individuals have conditions caused by outside factors?
b. What is the probability that three or more individuals have conditions caused by outside factors?
c. What are the mean and standard deviation of the number of individuals with conditions caused by outside factors?

Solution: Given:

Probability of heart failure due to outside factors: p= 13% or 0.13

Number of patients: n= 20

Let X be a random variable that represents the number of patients with heart failure due to outside factors. Therefore, X follows a binomial distribution i.e., $X \sim Binom(n = 20, p = 0.13)$. and the probability mass function (pmf) of X is given by

$$P(X = x) = 20Cx \times (0.13)^x \times (1 - 0.13)^{20-x};\ \text{x} = 0,1,2,\cdots,20$$

a. The probability that three individuals have conditions caused by outside factors:

$$P(X = 3) = 20C3 \cdot (0.13)^3 \cdot (1 - 0.13)^{20-3} \approx 0.2347$$

Interpretation: There is approximately a 23.47% chance that exactly 3 out of 20 patients will have heart failure caused by outside factors.

b. The probability that three or more individuals have conditions caused by outside factors: This can be calculated as:

$$P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

Calculate each term:

$$P(X = 0) = 20C0 \cdot (0.13)^0 \cdot (1 - 0.13)^{20-0} \approx 0.0617$$
$$P(X = 1) = 20C1 \cdot (0.13)^1 \cdot (1 - 0.13)^{20-1} \approx 0.1844$$
$$P(X = 2) = 20C2 \cdot (0.13)^2 \cdot (1 - 0.13)^{20-2} \approx 0.2618$$

Now add them up:

$$P(X \geq 3) = 1 - (0.0617 + 0.1844 + 0.2618) = 1 - 0.5079 \approx 0.4921$$

Interpretation: There is approximately a 49.21% chance that three or more individuals out of 20 will have heart failure caused by outside factors.

c. The mean and standard deviation of the number of individuals with conditions caused by outside factors: For a binomial distribution:

Mean , $\mu = np = 20 \times 0.13 = 2.6$

Standard deviation, $\sigma = \sqrt{np(1-p)} = \sqrt{20 \times 0.13 \times 0.87} = \sqrt{2.262} \approx 1.504$

Interpretation: On average, we expect about 2.6 patients out of 20 to have heart failure due to outside factors, with a standard deviation of approximately 1.504 patients.

Question 3. Suppose you have a fair spinner divided into four equal sections numbered 1 through 4, and a fair coin with two sides: heads (H) and tails (T). You spin the spinner and flip the coin simultaneously, and you record the number on the spinner and the result of the coin flip. Let's define a random variable Y as the number of times the spinner lands on 3 and the coin lands on heads (H) simultaneously. You conduct this experiment 15 times independently.

a. Write down the probability distribution function for this scenario.
b. Use the probability distribution function from (a) to determine the probability of getting the outcome (spinner lands on 3 and coin lands on H) "at least four times". Interpret your findings.
c. Determine the mean and standard deviation of this probability distribution function.

Solution:

a. Let us calculate the probability of getting a 3 on the spinner and heads (H) on the coin in a single trial.

Probability of spinner landing on 3, $P(3) = 0.25$

Probability of coin landing on heads (H), P(H) = 0.5

Therefore, probability of occurring both events (Spinner = 3 and Coin = H) in a single trial, $P(Y) = P(3) \times P(H) = 0.125$. Since the experiment was conducted 15 times independently, it follows binomial distribution with parameter, n = 15, p = 0.125. Thus, the probability mass function can be written as,

$$P(Y = y) = 15Cy\, 0.125^y\, (1 - 0.125)^{n-y}$$

b. $P(Y \geq 4) = 1 - P(Y < 4) = 1 - (P(0) + \cdots + P(3)) = 0.10778$

Comment: This probability 0.10778 represents the likelihood of getting the outcome of both the spinner landing on 3 and the coin landing on heads at least four times in 15 trials.

c. [Hints: Mean and Variance of binomial distribution]

Comment: On average, we can expect the spinner to land on 3 and the coin to land on heads about 1.875 times in 15 trials, with a standard deviation of approximately 1.28, which measures the variability of this outcome.

Question 4. A factory produces items and has a 12% defect rate. If an inspector randomly selects 15 items.

a. Calculate the probability of exactly 2 defective items. [Ans: 0.2870]
b. Calculate the mean number of defective items expected [Ans: 1.8]
c. Calculate the variance of the number of defective items. [Ans: 1.584]
d. Determine the probability of having at most 2 defective items. [Ans: 0.7346]

Question 5. At BRAC University, the probability that a randomly selected student is enrolled in the course STA201 is 0.4. Suppose 20 students are randomly chosen from the university.

a. What is the probability that exactly 8 out of these 20 students are enrolled in STA201? [Ans: 0.17971]
b. What is the probability that fewer than 2 out of these 20 students are enrolled in STA201? [Ans: 0.00519]
c. Calculate the mean and standard deviation of the binomial distribution representing the number of students enrolled in STA201 out of the 20 selected students. [Ans: Mean = 8, SD = 2.19]

Question 6. Based on data collected by the BracU authority, an estimate of the percentage of students who have at some point in their life been told they have hypertension is 23.53 percent. If we select a simple random sample of 20 students and assume that the probability that each has been told that he or she has hypertension is .24, find the probability that the number of students in the sample who have been told that they have hypertension will be:

a. Exactly three. [Ans: 0.14838]
b. Fewer than three. [Ans: 0.10855]
c. Between three and seven, inclusive. [Ans: 0.808]
d. How many students who have been told that they have hypertension would you expect to find in a sample of 20? [Ans: 4.8]

Question 7. The proportion of people with blood groups O, A, B and AB in a particular population are in the ratio 48:35:12:5, respectively. A random sample of 20 people from the population is picked.

a. What is the probability that exactly 10 with blood group O?
b. What is the probability that at most 2 with blood group AB?
c. Find the expected number of people with blood group A. [Ans: 7]

Solution:

Given that, $O: A: B: AB \ = \ 48: 35: 12: 5$

$P(O) \ = \ 48/(48 + 35 + 12 + 5) = \ 0.48$

P(A) = 0.35

P(B) = 0.12

P(AB) = 0.05

a. $P(O = 10) = \binom{20}{10}(0.48)^{10}(1 - 0.48)^{20-10} = 0.1734$
b. $P(AB \leq 2) = \binom{20}{0}(0.05)^{0} \cdot (1 - 0.05)^{20-0} + \binom{20}{1}(0.05)^{1} \cdot (1 - 0.05)^{20-1} = 0.92452$
c. Try yourself.

Question 8. In a particular section of the university, 70% of students are from the CSE department, others are from the BBA department. If 15 students are randomly selected then,

a. Write down the probability distribution function for this scenario. [$P(Y = y) = 15Cy\ (0.7)^{y}(0.3)^{n-y}$]
b. Calculate the mean number of students from CSE department. [Ans: 10.5]
c. Calculate the variance of the number from CSE department. [Ans: 3.15]
d. Determine the probability of having at exactly 12 Student from CSE department. [Ans: 0.17]

Question 9. An MCQ test has 10 questions each with 5 possible choices. A student did not study at all and chose all the answers randomly.

a. Probability that 2 answers will be correct. [n = 10, p = 1/5, Ans: 0.30199]
b. Probability that at least 1 answer will be correct. [Ans: 0.89263]
c. Probability that at most 2 answers will be correct. [Ans: 0.6778]
d. Probability that less than 1 answer will be correct. [Ans: 0.10737]
e. Average number of correct answers. [Ans: 2]
f. Variance of this distribution. [Ans: 1.6]

Question 10. 70% of the passengers who travel on the newly launched "Korean Coach" to Cox's Bazar from Dhaka buy 'Bangladesh Protidin' at the bookstall before they board the train. The train is full and each compartment holds 8 passengers.

a. What is the probability that all the passengers in a compartment have brought the newspaper? [Ans: 0.058]

b. What is the probability that none of the passengers in a compartment has brought the newspaper? [Ans: 6.561×10^{-5}]

c. What is the probability that at most two passengers have brought the newspaper? [Ans: 0.011]

d. Calculate expected number of passengers who have brought the newspaper. [Ans: 5.6]

Question 11. In a rural district of Bangladesh, a Hepatitis-B vaccination campaign was conducted to improve public health. It is known from previous health records that the Hepatitis-B inoculation is effective in 80% of cases. After 6 months, a health survey randomly selects 5 individuals from the vaccinated population to assess the effectiveness of the vaccination.

a. Determine the probability that exactly one out of the 5 selected individuals did not benefit from the Hepatitis-B inoculation.

b. Determine the probability that at least one of the 5 selected individuals did not benefit from the Hepatitis-B inoculation.

c. Find the expected number of individuals out of the 5 who did not benefit from the Hepatitis-B inoculation.

d. Calculate the variance and standard deviation of the number of individuals out of the 5 who did not benefit from the Hepatitis-B inoculation.

Solution:

Let X be the random variable representing the number of individuals out of 5 for whom the inoculation is not effective. Here, X follows a binomial distribution with:

Number of trials n=5

Probability of success (inoculation being effective) $p = 0.80$

Probability of failure (inoculation not being effective) $q = 1 - p = 0.20$

a. For k = 1, n = 5, and p= 0.80

$P(X = 1) = \binom{n}{k} p^{n-k} q^k \; = \; 0.409$

b. $P(X \geq 1) = 1 - P(x = 0) = 0.67$

c. $E(x) = n \cdot q = 1$

d. $V(x) = n \cdot p \cdot q$ **= 0.80**

Poisson Distribution

Question 1. In a city, the average number of traffic accidents at a specific intersection is 5 per day.

a. Determine the probability of having exactly 2 accidents on a given day.
b. Find the probability of having at least two accidents on a given day.
c. Find the probability of having at most one accident on a given day.
d. Find the probability of having no more than three accidents on a given day.

Solution: Given that the average number of traffic accidents at a specific intersection is 5 per day, we can model the situation using a Poisson distribution. Let X be the random variable representing the number of accidents on a given day. Since X follows a Poisson distribution with mean $\lambda = 5$, the probability mass function of X is given by:

$$P(X = x) = \frac{e^{-5}\,5^x}{x!}, x = 0,1,2,\cdots$$

a. Probability of having exactly 2 accidents on a given day:

$$P(X = 2) = \frac{e^{-5}\,5^2}{2!} \approx 0.0842$$

Interpretation: There is an 8.42% chance of exactly 2 accidents occurring on a given day.

b. Probability of having at least two accidents on a given day:

The probability of at least two accidents is the complement of having fewer than two accidents:

$$P(X \geq 2) = 1 - P(X < 2)$$

$$P(X < 2) = P(X = 0) + P(X = 1)$$

$$P(X = 0) = \frac{e^{-5}\,5^0}{0!} \approx 0.0067$$

$$P(X = 1) = \frac{e^{-1}\,5^1}{1!} \approx 0.0337$$

$$P(X \geq 2) = 1 - (0.0067 + 0.0337) \approx 0.9596$$

Interpretation: There is a 95.96% chance of having at least two accidents on a given day, indicating that it's highly likely more than one accident will occur.

c. Probability of having at most one accident on a given day:

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$P(X \leq 1) = 0.0067 + 0.0337 \approx 0.0404$$

Interpretation: There's only a 4.04% chance of having at most one accident, making it unlikely that a day will pass with one or fewer accidents.

d. Probability of having no more than three accidents on a given day:

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X = 3) = \frac{e^{-3}\, 5^3}{3!} \approx 0.1404$$

$$P(X \leq 3) = 0.0067 + 0.0337 + 0.0842 + 0.1404 \approx 0.2650$$

Interpretation: There is a 26.5% chance of having no more than three accidents on a given day, showing that it's somewhat likely but still less than half the time.

Question 2. A bookstore sells 27 novels in 9 hours. Assume the number of novel selves follows Poisson distribution.

a. Find the probability of selling exactly 5 novels in an hour.
b. What is the probability of selling 0 novels in an hour?
c. Calculate the mean and variance of novels sold per hour.
d. What is the probability of selling more than 4 novels in an hour?

Solution:

a. Find the probability of selling exactly 5 novels: $\lambda = 3, k = 5$

$$P(X = 5) = \frac{3^5 e^{-3}}{5!} = \frac{243e^{-3}}{120} \approx 0.10082$$

b. Probability of selling 0 novels:

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = e^{-3} \approx 0.0498$$

c. Mean and variance of novels sold: $\mu = \lambda = 3, \quad \sigma^2 = \lambda = 3$

d. Probability of selling more than 4 novels:

$$P(X > 4) = 1 - P(X \leq 4)$$

$$= 1 - \big(P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)\big) = 0.18474$$

Question 3. A library maintains records of daily visitor counts. On average, 10 people visit the library every hour. However, due to varying schedules and interests, the number of visitors can fluctuate. The library is open for 8 hours each day, and a study has shown that the number of visitors per hour follows a discrete distribution with a mean of 10 visitors per hour.

a. Write down the probability mass function of the distribution.
b. What is the probability that exactly 8 people visit the library in a given hour?
c. What is the probability that at least 5 people visit the library in a given hour?
d. What is the probability that the library gets fewer than 4 visitors in a two-hour period?
e. What is the average number of visitors per hour and the standard deviation of the number of visitors per hour?

Solution:

a. Let, X = Number of visitors per hour, and it follows Poisson distribution with parameter $\lambda = 10$ Thus, the distribution function can be written as:

$$P(X) = \frac{e^{-10}10^{x}}{x!}$$

b. $P(X = 8) = 0.1126$

Comment: The probability that exactly 8 people visit the library in a given hour is approximately 0.1126.

c. $P(X \geq 5) = 1 - P(X < 5) = 0.9707$

Comment: The probability that at least 5 people visit the library in a given hour is approximately 0.9622.

d. Here, $\lambda = 2 \times 20$; [Two-hour period]

$$P(X < 4) = P(0) + \cdots + P(3) = 0.0000032$$

Comment: The probability that the library gets fewer than 4 visitors in a two-hour period is approximately 0.0000032.

e. Try yourself.

Question 4. During a recent awareness campaign in Dhaka, the emergency response center received an average of 8 emergency calls per hour related to the event.

a. What is the probability that the center will receive exactly 7 emergency calls in an hour? [Ans: 0.1396]
b. What is the probability that the center will receive at most 5 emergency calls in an hour? [Ans: 0.1912]
c. What is the probability that the center will receive more than 6 emergency calls in an hour? [Ans: 0.6866]
d. What is the probability that the center will receive exactly 24 emergency calls in a span of 4 hours? [Ans: 0.03 (when $\lambda = 32$)]

Question 5. The number of serious gastrointestinal reactions reported to the British Committee on Safety of Medicine was 538 for 9,160,000 prescriptions of the anti-inflammatory drug piroxicam. This corresponds to a rate of .06 gastrointestinal reactions per 1000 prescriptions written. Using a Poisson model for probability, with $\lambda = 0.06$, find the probability of

a. Exactly one gastrointestinal reaction in 1000 prescriptions. [Ans: 0.05651]
b. No gastrointestinal reactions in 1000 prescriptions. [Ans: 0.94176]
c. Exactly two gastrointestinal reactions in 1000 prescriptions. [Ans: 0.00170]
d. At least one gastrointestinal reaction in 2000 prescriptions. [Ans: 0.11308 (when $\lambda = 0.12$)]

Question 6. A private clinic monitors the number of emergency patients it receives daily. Over time, it has been observed that the clinic averages 10 emergency patients per day. The arrival of these patients is random and follows a pattern that can be modeled using the Poisson distribution, which is commonly used to predict the probability of a given number of events happening within a fixed interval of time when the events occur independently.

a. What is the probability that the clinic will receive exactly 5 emergency patients on a specific day? [Ans: 0.03783]
b. What is the probability that the clinic will receive at least 3 emergency patients on a specific day? [Ans: 0.997]
c. What is the probability that the clinic will receive between 5 and 10 emergency patients (inclusive) on a specific day? [Ans: 0.554]

Question 7. A group of BRAC University students created a ride-share app where customers and riders can negotiate fares. Their helpline center received an average of 6 calls per hour related to software bugs.

a. Write down the probability mass function of the distribution. $\left[P(X) = \frac{e^{-6}6^x}{x!}\right]$

b. What is the probability that they will receive exactly 8 calls per hour? [Ans: 0.10]

c. What is the probability that they will receive at most 4 calls per hour? [Ans: 0.2851]

d. What is the probability that they will receive more than 3 calls per hour? [Ans: 0.8488]

Question 8. On average, a factory experiences 2 machine failures per day.

a. What is the probability that the factory will have no machine failures on a given day? [Ans: 0.1353]

b. Find the probability of having at least three machine failures on a given day. [Ans: 0.32332]

c. Find the probability of having at most two machine failures on a given day. [Ans: 0.67668]

d. Determine the probability of having exactly 5 machine failures on a given day. [Ans: 0.03609]

Question 9.

According to the National Office of Vital Statistics of the Vancouver Health and Human Services, the average number of accidental drownings per year in the state is 3.0 per 10,0000 populations. Find the probability that in a city of population 20,0000, there will be-

a. 2 accidental drownings per year. [Ans: 0.0446]

b. At least 3 accidental drowning per year. [Ans: 0.938]

c. Between 4 and 8 accidental drowning per year. [Ans: 0.6964]

Solution:

Let, X be the random variable representing number of accidental drownings per year

Here, Average number of accidental drownings per year, $\lambda = 200000 \times \frac{3}{100000} = 6$

So, $X \sim Poisson(\lambda = 6)$

a. Try yourself.

b. Try yourself.

c. Try yourself.

Question 10. In a vintage vinyl record factory in Bangladesh, the quality control team has observed that the number of defects per record follows a Poisson distribution with a mean of 0.5 defects per record. The factory produces limited edition records that are highly valued by collectors. The management aims to ensure that each record meets high-quality standards before being sold.

a. Determine the probability that a randomly selected record will have exactly two defects. [Ans: 0.0758]

b. Given that the factory wants to ensure all records sold to customers are defect-free, calculate the percentage of records that will be rejected for sale because they have defects. **(Hint: Use $e^{-0.5} \approx 0.6065$ for your calculations.)** [Ans: 39.35%]

c. Suppose the factory wants to set a new quality standard where only records with at most one defect are allowed for sale. Calculate the probability that a randomly selected record will meet this new standard. [Ans: 0.9098]

d. Calculate the variance and standard deviation of the number of defects per record in the factory's production.

Normal Distribution

Question 1. Tool workers are subject to work-related injuries. One disorder, caused by strains to the hands and wrists, is called carpal tunnel syndrome. It strikes as many as 23,000 workers per year. The U.S. Labor Department estimates that the average cost of this disorder to employers and insurers is approximately $30,000 per injured worker. Suppose these costs are normally distributed, with a standard deviation of $9,000.

a. What proportion of the costs are between $15,000 and $45,000?
b. What proportion of the costs are greater than $50,000?
c. What proportion of the costs are between $5,000 and $20,000?
d. Suppose the standard deviation is unknown, but 90.82% of the costs are more than $7,000. What would be the value of the standard deviation?
e. Suppose the mean value is unknown, but the standard deviation is still $9,000. How much would the average cost be if 79.95% of the costs were less than $33,000?

Solution: Given that, Mean cost, $\mu =$ \$30,000; Standard deviation $\sigma =$ \$9,000

Let X be a random variable representing the costs. Therefore, X follows a normal distribution with mean, $\mu =$ \$30,000 and standard deviation $\sigma =$ \$9,000.

a. Proportion of the costs are between $15,000 and $45,000:

$$P(15000 < X < 45000) = P\left(\frac{15000-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{45000-\mu}{\sigma}\right)$$

$$= P\left(\frac{15000-30000}{9000} < Z < \frac{45000-30000}{9000}\right); \left[\because X \sim N(\mu, \sigma^2),\ Z = \frac{X-\mu}{\sigma} \sim N(0,1)\right]$$

$$= P(-1.67 < Z < 1.67)$$

$$= \mathrm{P}(\mathrm{Z} < 1.67) - \mathrm{P}(\mathrm{Z} < -1.6)$$

$$= 0.9525 - 0.0475 = 0.9050;\ [\text{Using standard normal distribution tables}]$$

So, approximately 90.5% of the costs are between $15,000 and $45,000.

b. Proportion of the costs are greater than $50,000:

$$P(X > 50000) = P\left(\frac{X-\mu}{\sigma} > \frac{50000-\mu}{\sigma}\right)$$

$$= P\left(Z > \frac{50000-30000}{9000}\right)$$

$$= P(Z > 2.22)$$

$$= 1 - P(Z < 2.22)$$

$$= 1 - 0.9868 = 0.0132$$

So, approximately 1.32% of the costs are greater than $50,000.

c. Proportion of the costs those are between \$5,000 and \$20,000:

$$P(5000 < X < 20000) = P\left(\frac{5000-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{20000-\mu}{\sigma}\right)$$

$$= P\left(\frac{5000-30000}{9000} < Z < \frac{20000-30000}{9000}\right)$$

$$= P(-2.78 < Z < -1.11)$$

$$= P(Z < -1.11) - P(Z < -2.78)$$

$$= 0.1335 - 0.0027 = 0.1308$$

So, approximately 13.08% of the costs are between \$5,000 and \$20,000.

d. Suppose the standard deviation is unknown, but 90.82% of the costs are more than \$7,000. What would be the value of the standard deviation?

We are given $P(X > 7{,}000) = 0.9082$, which means

$$P(X \leq 7{,}000) = 1 - 0.9082 = 0.0918$$

$$or, \quad P\left(\frac{X-\mu}{\sigma} \leq \frac{7{,}000-\mu}{\sigma}\right) = 0.0918$$

$$or, \quad P\left(Z \leq \frac{7{,}000-30000}{\sigma}\right) = 0.0918$$

Find the z-score corresponding to 0.0918 in the standard normal distribution, which is approximately -1.33.

Therefore,

$$\frac{7{,}000 - 30{,}000}{\sigma} = -1.33$$

$$or, \quad \sigma = \frac{30{,}000 - 7{,}000}{1.33} = \frac{23{,}000}{1.33} \approx 17{,}293.23$$

So, the standard deviation would be approximately \$17,293.23.

e. Suppose the mean value is unknown, but the standard deviation is still \$9,000. How much would the average cost be if 79.95% of the costs were less than \$33,000?

We are given,

$$P(X \leq 33{,}000) = 0.7995$$

$$or, \quad P\left(\frac{X-\mu}{\sigma} \leq \frac{7{,}000-\mu}{\sigma}\right) = 0.7995$$

$$or, P\left(Z \leq \frac{7{,}000-\mu}{9000}\right) = 0.7995$$

Find the z-score corresponding to 0.7995 in the standard normal distribution, which is approximately 0.84. Therefore,

$$\frac{33{,}000 - \mu}{9{,}000} = 0.84$$

$$or, \quad \mu = 33{,}000 - (0.84 \times 9{,}000) = 33{,}000 - 7{,}560 = 25{,}440$$

So, the average cost would be approximately \$25,440.

Question 2. A bank is analyzing the transaction amounts of its customers. The transactions are normally distributed with a mean of 70 and a standard deviation of 3.

a. What is the probability that a randomly selected transaction is more than \$75?
b. What is the probability that a randomly selected transaction is between \$68 and \$73? What amount corresponds to the top 10% of all transactions?
c. If we randomly sample 25 customers, what is the probability that their average transactions will be greater than \$71?

Solution:

a. Probability that a randomly selected transaction is more than \$75:

$$Z = \frac{75 - 70}{3} \approx 1.67$$

Using Z-tables,

$$P(Z > 1.67) \approx 1 - 0.9525 = 0.0475$$

b. Percentage of transaction between \$68 and \$73:

$$Z_1 = \frac{68 - 70}{3} = -0.67, \quad Z_2 = \frac{73 - 70}{3} = 1$$

Using Z-tables:

$$P(68 < X < 73) = P(Z < 1) - P(Z < -0.67) \approx 0.8413 - 0.2514 \approx 0.5899$$

Transaction corresponding to the top 10% means 90th percentile: From Z-tables, (Z),

$$X = \mu + Z\sigma = 70 + (1.28)(3) \approx 73.84$$

c. Probability that averages transaction of 25 customer is greater than \$71:

$$Z = \frac{71 - 70}{\frac{\sigma}{\sqrt{n}}} = \frac{1}{\frac{3}{\sqrt{25}}} = \frac{1}{0.6} \approx 1.67$$

Using Z-tables: $P(Z > 1.67) \approx 0.0475$

Question 3. In the aftermath of the recent flood in Bangladesh, the depth of water (in inches) measured in 300 different locations was found to be normally distributed with a mean of 68.0 inches and a standard deviation of 3.0 inches. Based on this information:

a. How many locations had a water depth of at least 65 inches? [Ans: Number of locations: 0.8413×300≈252]
b. How many locations had a water depth of at most 80 inches? [Ans: 300]
c. How many locations had a water depth between 60 inches and 65 inches? [Ans: 46]
d. How many locations had a water depth equal to 70 inches? [Ans: 0]

Question 4. If the total cholesterol values BracU faculty members are approximately normally distributed with a mean of 200 mg/100 ml and a standard deviation of 20 mg/100 ml, find the probability that an individual picked at random from this population will have a cholesterol value:

a. between 180 and 200 mg/100ml. [Ans: 0.3413]
b. greater than 225 mg/100ml. [Ans: 0.1056]
c. less than 150 mg/100ml. [Ans: 0.00621]
d. between 190 and 210 mg/100ml. [Same as (a)]

Question 5. The Phillips Bangladesh manufactures electric bulbs that have a length of life that is normally distributed with mean equal to 800 hours and standard deviation of 40 hours.

a. Find the probability that a bulb burns between 778 and 834 hours. [Ans: 0.5111]
b. Find the probability that a bulb burns after 834 hours. [Ans: 0.1977]

Question 6. The monthly salaries of newly graduated students from BRAC University are normally distributed, with a mean of $250 and a standard deviation of $6. Based on this information find the probability that a randomly selected graduated will get salary:

a. At least $235. [Ans: 0.993]
b. Between $255 to $260. [Ans: 0.16]
c. At most $300. [Ans: 1]
d. Between $235 to $280. **[Hint: similar to (b)]**

Question 7. Assume that, X is defined as the GPA score of students of the department of CSE of BRAC university in their term final examination was found to follow a normal distribution with a mean of 2.1 and a standard deviation of 0.6.

a. Find the probability that a given student will score at most 2. [Ans: 0.4325]
b. Find the probability that a given student will score between 2.5 and 3.5. [Ans: 0.2416]
c. Find the probability that a given student will score at least 3. [Ans: 0.0668]

Question 8. A recent study of the hourly wages of maintenance crews for major airlines showed that the mean hourly salary was 165$ with standard deviation of 35$. If we select a crew member at random

a. What is the probability that the crew member earns between 180$ to 200$ per hour? [Ans: 0.1749]
b. What is the probability that the crew member earns at least 125$ per hour? [Ans: 0.8729]
c. In what interval would you expect the central 95% of the value to be found? [Hints: 3sigma rules. Ans: 95, 235]

Question 9. An expert witness in a paternity suit testifies that the length (in days) of human gestation is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short gestation indicated by the testimony?

Solution:

Let X denote the length of the gestation, and assume that the defendant is the father. Then the probability that the birth could occur within the indicate period is

$P(X > 290 \mid X < 240)$

$= P(X > 290) + P(X < 240)$

$= P\left\{\frac{X - 270}{10} > 2\right\} + P\left\{\frac{X - 270}{10} < -3\right\}$

$= 1 - \Phi(2) + 1 - \Phi(3) \approx .0241$

Geometric Distribution

Question 1. The probability of a successful optical alignment in the assembly of an optical data storage product is 0.8. Assume the trials are independent.

a. What is the probability that the first successful alignment requires exactly four trials?
b. What is the probability that the first successful alignment requires at most four trials?
c. What is the probability that the first successful alignment requires at least four trials?

Solution: Consider the trials are independent and the probability of success (successful optical alignment) is $p = 0.8$. We define X as the number of trials until the first success, so X follows a geometric distribution with $p = 0.8$. i.e., $X \sim Geom(0.8)$ and the probability mass function (pmf) of X is given by: $P(X = x) = 0.8 \times (1 - 0.8)^{x-1}\ for\ x = 1,2,3, \ldots$

a. The probability that the first successful alignment requires exactly four trials:

   For $x = 4$: $P(X = 4) = 0.8 \times (1 - 0.8)^{4-1} = 0.0064$

   So, the probability that the first successful alignment requires exactly four trials is 0.0064.

b. The probability that the first successful alignment requires at most four trials:

   This is the cumulative probability of the first success occurring within the first four trials:

$$P(X \leq 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

   Calculating each:

$$P(X = 1) = 0.8 \times (0.2)^0 = 0.8$$
$$P(X = 2) = 0.8 \times (0.2)^1 = 0.16$$
$$P(X = 3) = 0.8 \times (0.2)^2 = 0.032$$
$$P(X = 4) = 0.8 \times (0.2)^3 = 0.0064$$

   Adding these together:

$$P(X \leq 4) = 0.8 + 0.16 + 0.032 + 0.0064 = 0.9984$$

   So, the probability that the first successful alignment requires at most four trials is 0.9984.

c. The probability that the first successful alignment requires at least four trials. This is the complementary probability of the first success happening after three trials:

$$P(X \geq 4) = 1 - P(X \leq 3)$$

   From part (b): $P(X \leq 3) = 0.8 + 0.16 + 0.032 = 0.992$

   Therefore: $P(X \geq 4) = 1 - 0.992 = 0.008$

   So, the probability that the first successful alignment requires at least four trials is 0.008.

Question 2. In the programming club at BRAC University, members occasionally encounter bugs while working on projects. It is known from previous experience that the probability of encountering a bug during any coding session is 3%. To evaluate their debugging process, the club decides to monitor the occurrence of bugs during their coding sessions.

a. What is the probability that the first bug is encountered during the 6th coding session?
b. What is the probability that the first bug is found within the first 3 coding sessions?
c. On average, how many sessions would you expect to conduct before encountering the first bug?
d. Calculate the variance and standard deviation of the number of sessions until the first bug is found.
e. Calculate the coefficient of variation (CV) for the number of sessions until the first bug is found.

Solution: Let X be the number of coding sessions until the first bug is encountered. Since the probability of encountering a bug during any coding session is $p = 0.03$, X follows a geometric distribution with parameter p.

a. The probability that the first bug is encountered during the 6^{th} session is given by:

$$P(X = 6) = (1-p)^{6-1} \times p = (1-0.03)^5 \times 0.03 = 0.0251$$

So, the probability is 0.0251.

b. The probability that the first bug is found within the first 3 sessions is:

$$P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3)$$
$$P(X = 1) = 0.03$$
$$P(X = 2) = 0.97 \times 0.03 = 0.0291$$
$$P(X = 3) = (0.97)^2 \times 0.03 \approx 0.941 \times 0.03 \approx 0.0274$$
$$P(X \leq 3) = 0.03 + 0.0291 + 0.0282 = 0.0873$$

So, the probability is 0.0873.

c. The expected value (mean) of X for a geometric distribution is:

$$E(X) = \frac{1}{p} = \frac{1}{0.03} \approx 33.33$$

So, on average, you would expect to conduct 33.33 sessions before encountering the first bug.

d. The variance of X for a geometric distribution is:

$$Var(X) = \frac{1-p}{p^2} = \frac{1-0.03}{(0.03)^2} = \frac{0.97}{0.0009} \approx 1077.78$$

The standard deviation is the square root of the variance:

$$SD(X) = \sqrt{1077.78} \approx 32.82$$

e. The coefficient of variation (CV) is given by:

$$CV = \frac{SD(X)}{E(X)} = \frac{32.82}{33.33} \approx 0.9847$$

So, the coefficient of variation is approximately 0.9847.

Question 3. Max decides to implement a quality control process to minimize the number of defective bulbs reaching customers. He instructs his team to inspect every batch of 75 bulbs and remove any defective ones before packaging. Unfortunately, he found that, on average, there were 3 defective bulbs in every batch.

a. What is the probability that Max will find the first faulty lightbulb on the 6th one that he tested? [Ans: 0.0326]

b. What is the probability that it takes at least six trials until he finds the first defective lightbulb?

c. Find the number of lightbulbs we would expect Max to inspect until he finds his defective, as well as the standard deviation? [Mean = 25, SD = 24.5]

Question 4. A die is thrown until 6 occurs for the first time. What is the probability that

a. Exactly 3 tosses will be required. [Ans: 0.69]

b. More than two tosses will be required.

c. Find the average number of tosses required. [Ans: 6]

d. Find the variance of this distribution.

Question 5. The probability that a player qualifies in a national chess championship is 0.7.

a. Find the probability that a player chosen at random qualifies the championship on the third attempt. [Ans: 0.063]

b. Find the probability that a player chosen at random qualifies the championship before the fourth attempt. [Ans: 0.973]

c. Determine the mean and standard deviation of the above distribution. [Mean = 1.428, SD = 0.78]