(2) Event B = The selected candidate scores more than 60 marks

$P(B)$ = relative frequency for candidates scoring more than 60 marks.

$$= \frac{\text{No. of candidates scoring more than 60 marks}}{\text{Total number of candidates in the sample}} = \frac{m}{n}$$

m = No. of candidates scoring more than 60 marks

$= 326 + 124$

$= 450$

Now, $P(B) = \frac{m}{n}$

$= \frac{450}{1191}$

$= \frac{150}{397}$

Required probability $= \frac{150}{397}$

(3) Event C = The selected candidate scores from 21 to 80 marks

$P(C)$ = relative frequency for candidates scoring from 21 to 80 marks.

$$= \frac{m}{n} = \frac{\text{No. of candidates scoring from 21 to 80 marks}}{\text{Total number of candidates in the sample}}$$

m = No. of candidates scoring from 21 to 80 marks

$= 162 + 496 + 326$

$= 984$

Now, $P(C) = \frac{m}{n}$

$= \frac{984}{1191}$

$= \frac{328}{1191}$

Required probability $= \frac{328}{1191}$

Illustration 37 : A factory runs in two shifts. The sample data about the quality of items produced in these shifts are shown in the following table :

Quality	**Shift**		**Total**
	I	**II**	
Defective items	24	46	70
Non-defective items	2176	2754	4930
Total	**2200**	**2800**	**5000**

One item is randomly selected from the production of the factory.

(1) If the item is taken from the production of the first shift then find the probability that it is defective.

(2) If the item is defective then find the probability that it is taken from the production of the first shift.

The total number of units in the sample = 5000

We shall define the events as follows :

Event A = The selected item is from the production of first shift

$$P(A) = \frac{\text{No. of items produced in the first shift}}{\text{Total number of items in the sample}} = \frac{m}{n}$$

m = No. of items produced in the first shift

= 2200

Now, $P(A) = \frac{m}{n}$

$= \frac{2200}{5000}$

Event D = The selected item is defective

$P(D)$ = relative frequency for defective items

$$= \frac{\text{No. of defective items}}{\text{Total number of items in the sample}} = \frac{m}{n}$$

m = No. of defective items

= 70

Now, $P(D) = \frac{m}{n}$

$= \frac{70}{5000}$

Event $A \cap D$ = The selected item is produced in the first shift and it is defective

$P(A \cap D)$ = relative frequency for event $A \cap D$

$$= \frac{\text{No. of items in event } A \cap D}{\text{Total number of items in the sample}} = \frac{m}{n}$$

m = No. of items in event $A \cap D$

= 24

Now, $P(A \cap D) = \frac{m}{n} = \frac{24}{5000}$

(1) The event that the item is defective when it is taken from the first shift = D/A

Probability of D/A using the formula of conditional event

$$P(D/A) = \frac{P(A \cap D)}{P(A)}$$

$$= \frac{\frac{24}{5000}}{\frac{2200}{5000}}$$

$$= \frac{24}{2200}$$

$$= \frac{3}{275}$$

Required probability = $\frac{3}{275}$

(This probability can be directly obtained as relative frequency $\frac{24}{2200}$ of the event D/A.)

(2) The event that the item is taken from the first shift when it is defective = A/D

Probability of A/D using the formula of condition probability

$$P(A/D) = \frac{P(A \cap D)}{P(D)}$$

$$= \frac{\frac{24}{5000}}{\frac{70}{5000}}$$

$$= \frac{24}{70}$$

$$= \frac{12}{35}$$

Required probability $= \frac{12}{35}$

(This probability can be directly obtained as relative frequency $\frac{24}{70}$ of the event A/D.)

Limitations : The limitations of the statistical definition of probability are as follows :

(1) The value of probability can be obtained by the statistical definition of probability only if $n \to \infty$ that is if n tends to infinity. But the infinite value of n can not be taken in practice.

(2) The probability obtained by this definition is an estimated value. The exact value of probability cannot be known using this definition.

Exercise 1.5

1. The sample data about monthly travel expense (in ₹) of a large group of travellers of local bus in a megacity are given in the following table :

Monthly travel expense (₹)	501–600	601–700	701–800	801–900	901 or more
No. of travellers	318	432	639	579	174

One person from this megacity travelling by local bus is randomly selected. Find the probability that the monthly travel expense of this person will be (1) more than ₹ 900 (2) at the most ₹ 700 (3) ₹ 601 or more but ₹ 900 or less.

2. The details of a sample inquiry of 4979 voters of constituency are as follows :

Details	**Males**	**Females**
Supporters of party *A*	1319	1118
Supporters of party *B*	1217	1325

One voter is randomly selected from this constituency.

(1) If this voter is a male, find the probability that he is a supporter of Party A.

(2) If this voter is a supporter of Party A, find the probability that he is a male.

*

Summary

- The events based on chance are called random events.
- The experiment which can be independently repeated under identical conditions and all its posible outcomes are known but which of the outcomes will appear can not be predicted with certainty before conducting the experiment is called a random experiment.
- The set of all possible outcomes of a random experiment is called a sample space of that experiment.
- A subset of the sample space of random experiment is called an event of that random experiment.
- U is a finite sample space and A and B are two events in it. If events A and B can never occur together that is if $A \cap B = \phi$ then the events A and B are called mutually exclusive events.
- If the group of favourable outcomes of events of a random experiment is the sample space then the events are called exhaustive events.
- The elementary events are mutually exclusive and exhaustive.
- If there is no apparent reason to believe that out of one or more events of a random experiment, any one event is more or less likely to occur than the other events then the events are called equi-probable events.
- The number of mutually exclusive, exhaustive and equi-probable outcomes in the sample space U of a random experiment is n. If m outcomes among them are favourable for the event A then probability of event A is $\frac{m}{n}$.
- The range of values of $P(A)$, the probability of any event A of a sample space U, is 0 to 1. That is $0 \le P(A) \le 1$.
- A and B are any two events in a finite sample space U. If the probability of occurrence of event A does not change due to occurrence (or non-occurrence) of event B then A and B are independent events.
- Suppose a random experiment is repeated n times under identical conditions. If an event A occurs in m trials out of n trials then the relative frequency $\frac{m}{n}$ of event A gives the estimate of the probability of event A, $P(A)$. When the larger and larger value of n is taken that is when n tends to infinity, the limiting value of $\frac{m}{n}$ is called the probability of event A.

List of Formulae

(1) Complementary event of A $\quad A' = U - A$

(2) Difference event of A and B $\quad A - B = A \cap B' = A - (A \cap B)$ (only event A occurrs.)

(3) Difference event of B and A $\quad B - A = A' \cap B = B - (A \cap B)$ (only event B occurrs.)

(4) The probability of an event A of the sample space of a random experiment is $P(A) = \frac{m}{n}$.

(5) Law of addition of probability

For two events A and B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

For any three events A, B and C,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

If two events A and B are mutually exclusive,

$$P(A \cup B) = P(A) + P(B)$$

If three events A, B and C are mutually exclusive,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

If two events A and B are mutually exclusive and exhaustive,

$$P(A \cup B) = P(A) + P(B) = 1$$

If three events A, B and C are mutually exclusive and exhaustive,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) = 1$$

(6) Conditional probability

Event B occurs under the condition that event A occurs

$$P(B/A) = \frac{P(A \cap B)}{P(A)}; \quad P(A) \neq 0$$

Event A occurs under the condition that event B occurs

$$P(A/B) = \frac{P(A \cap B)}{P(B)}; \quad P(B) \neq 0$$

(7) Law of multiplication of probability

For any two events A and B,

$P(A \cap B) = P(A) \times P(B/A)$; $P(A) \neq 0$

$P(A \cap B) = P(B) \times P(A/B)$; $P(B) \neq 0$

- For independent events A and B,

$P(A \cap B) = P(A) \times P(B)$

$P(A' \cap B') = P(A') \times P(B')$

$P(A' \cap B) = P(A') \times P(B)$

$P(A \cap B') = P(A) \times P(B')$

(8) According to statistical definition of probability,

$P(A) = \lim_{n \to \infty} \frac{m}{n}$

Exercise 1

Section A

Find the correct option for the following multiple choice questions :

1. Which event is given by a special subset ϕ of the sample space U ?

(a) Certain event (b) Complementary event of ϕ

(C) Union of events U and ϕ (d) Impossible event

2. What is the value of $P(A \cap A')$ for events A and A' ?

(a) 1 (b) 0 (c) 0.5 (d) between 0 and 1

3. Which of the following options is true for any event of the sample space ?

(a) $P(A) < 0$ (b) $0 \leq P(A) \geq 1$ (c) $0 \leq P(A) \leq 1$ (d) $P(A) > 1$

4. Which of the following options is not true for any two events A and B in the sample space U where $A \subset B$?

(a) $P(A \cap B) = P(B)$ (b) $P(A \cap B) = P(A)$

(c) $P(A \cup B) \geq P(A)$ (d) $P(B - A) = P(B) - P(A)$

5. What is the other name of the classical definition of probability ?

(a) Mathematical definition (b) Axiomatic definition

(c) Statistical definition (d) Geometric definition

6. Which of the following statement for probability of elementary events H and T of random experiment of tossing a balanced coin is not true ?

(a) $P(T)=0.5$ (b) $P(H)+P(T)=1$ (c) $P(H \cap T)=0.5$ (d) $P(H)=0.5$

7. Which random experiment from the following random experiments has an infinite sample space ?

(a) Throwing two dice (b) Selecting two employees from an office

(c) To measure the life of electric bulb (d) Select a card from 52 cards

8. If $A \cup A' = U$ then what type of events are A and A' ?

(a) Independent events (b) Complementary events

(c) Certain events (d) Impossible events

9. If $P(A/B) = P(A)$ and $P(B/A) = P(B)$ then what type of events are A and B ?

(a) Independent events (b) Complementary events

(c) Certain events (d) Impossible events

10. Two events A and B of a sample space are mutually exclusive. Which of the following will be equal to $P(B-A)$?

(a) $P(A)$ (b) $P(B)$ (c) $P(A \cap B)$ (d) $P(A \cup B)$

11. What is the total number of sample points in the sample space formed by throwing three six-faced balanced dice simultaneously ?

(a) 6^2 (b) 3^6 (c) 6×3 (d) 6^3

12. If one number is randomly selected between 1 and 20, what is the probability that the number is a multiple of 5 ?

(a) $\frac{1}{2}$ (b) $\frac{1}{6}$ (c) $\frac{1}{5}$ (d) $\frac{1}{3}$

13. If events A and B are independent, which of the following options is true ?

(a) $P(A \cap B) = P(A) \times P(B)$ (b) $P(A \cup B) = P(A) + P(B)$

(c) $P(A \cup B) = P(A) \times P(B)$ (d) $P(A \cap B) = P(A) + P(B)$

14. What is the probability of having 5 Thursdays in the month of February in a year which is not a leap year ?

(a) 0 (b) $\frac{1}{7}$ (c) $\frac{2}{7}$ (d) $\frac{3}{7}$

15. If $P(A)=0.4$ and $P(B')=0.3$ for two independent events A and B of a sample space then state the value of $P(A \cap B)$.

(a) 0.12 (b) 0.42 (c) 0.28 (d) 0.18

16. For two events A and B of a samples space, state the event $(A \cap B) \cup (A \cap B')$.

(a) ϕ (b) B (c) A (d) U

17. According to the mathematical definition of probability, what is the probability of each outcome among the n outcomes of a random experiment ?

(a) 0 (b) $\frac{1}{n}$ (c) 1 (d) can not say

Section B

Answer the following questions in one sentence :

1. Give two examples of random experiment.

2. Draw the Venn diagram for A-B, the difference event of A and B.

3. Define an event.

4. Write the sample space of a random experiment of throwing one balanced die and a balanced coin simultaneously.

5. Define conditional probability.

6. State the formula for the probability of occurrence of at least one event out of three events A, B and C.

7. Define independent events.

8. Write the law of multiplication of probability for two independent events A and B in a sample space.

9. Interpret $P(A/B)$ and $P(B/A)$.

10. When can we say that three events A, B and C in a sample space are exhaustive ?

11. Arrange $P(A \cup B), P(A), P(A \cap B), 0, P(A)+P(B)$ in the ascending order.

12. Define :

(1) Random Experiment (2) Sample Space

(3) Equi-probable Events (4) Favourable Outcomes

(5) Probability (Mathematical definition) (6) Probability (Statistical definition)

(7) Impossible Event (8) Certain Event

13. For two events A and B in a sample space, $A \cap B = \phi$ and $A \cup B = U$. State the values of $P(A \cap B)$ and $P(A \cup B)$.

14. If two events A and B in a sample space are independent then state the formula for $P(A \cup B)$.

15. If $A = \{x \mid 0 < x < 1\}$ and $B = \{x \mid \frac{1}{4} \le x \le 3\}$ then find $A \cap B$.

16. For two independent events A and B, $P(A) = 0.5$ and $P(B) = 0.7$. Find $P(A' \cap B')$.

17. If $P(A) = 0.8$ and $P(A \cap B) = 0.25$, find $P(A - B)$.

18. If $P(A) = 0.3$ and $P(A \cap B) = 0.03$, find $P(B/A)$.

19. If $P(A) = P(B) = K$ for two mutually exclusive events A and B, find $P(A \cup B)$.

20. If $P(A' \cap B) = 0.45$ and $A \cap B = \phi$, find $P(B)$.

21. Two events A and B in a sample space are mutually exclusive and exhaustive. If $P(A) = \frac{1}{3}$, find $P(B)$.

22. 2% items in a lot are defective. What is the probability that an item randomly selected from this lot is non-defective ?

23. State the number of sample points in the random experiment of tossing five balanced coins.

24. State the number of sample points in the random experiment of tossing one balanced coin and two balanced dice simultaneously.

25. Is it possible that $P(A) = 0.7$ and $P(A \cup B) = 0.45$ for two events A and B in a sample space ?

26. Two cards are selected one by one with replacement from 52 cards. State the number of elements in the sample space of this random experiment.

27. For two independent events A and B, $P(B/A) = \frac{1}{2}$ and $P(A \cap B) = \frac{1}{5}$. Find $P(A)$.

28. 1998 tickets out of 2000 tickets do not have a prize. If a person randomly selects one ticket from 2000 tickets then what is the probability that the ticket selected is eligible for prize ?

Section C

Answer the following questions :

1. Define the following events and draw their venn diagram :

(1) Mutually exclusive events
(2) Union of events
(3) Intersection of events
(4) Difference event
(5) Exhaustive events
(6) Complementary event

2. Give the illustrations of finite and infinite sample space.

3. Give the illustrations of impossible and certain event.

4. State the characteristics of random experiment.

5. State the assumptions of mathematical definition of probability.

6. State the limitations of mathematical definition of probability.

7. State the limitations of statistical definition of probability.

8. Explain the equiprobable events with illustration.

9. State the law of addition of probability for two events A and B. Write the law of addition of probability if these two events are mutually exclusive.

10. State the law of multiplication of probability for two events A and B. Write the law of multiplication of probability if these two events are independent.

11. State the following results for two independent events A and B :

(1) $P(A \cap B)$
(2) $P(A' \cap B')$
(3) $P(A \cap B')$
(4) $P(A' \cap B)$

12. If $P(A) = \frac{1}{3}$, $P(B) = \frac{2}{3}$ and $P(A \cap B) = \frac{1}{6}$ then find $P(A' \cap B')$.

13. If $P(B) = 2P(A/B) = 0.4$ then find $P(A \cap B)$.

14. If the events A and B are independent and $3P(A) = 2P(B) = 0.12$ then find $P(A \cap B)$.

15. If $5P(A) = 3P(B) = 2\,P(A \cup B) = \frac{3}{2}$ for two events A and B then find $P(A' \cup B')$.

16. If $P(A \cap B) = 0.12$ and $P(B) = 0.3$ for two independent events A and B then find $P(A \cup B)$.

17. If $A = \{x \mid 1 < x < 3\}$ and $B = \{x \mid \frac{1}{2} \leq x < 2\}$ then find $A \cup B$ and $A \cap B$.

18. The probability of occurrence of at least one of the two events A and B is $\frac{1}{4}$. The probability that event A occurs but event B does not occur is $\frac{1}{5}$. Find the probability of event B.

19. If $P(B) = \frac{3}{5}$ and $P(A' \cap B) = \frac{1}{2}$, for two events A and B, find $P(A/B)$.

20. 6 persons have a passport in a group of 10 persons. If 3 persons are randomly selected from this group, find the probability that

(1) all the three persons have a passport

(2) two persons among them do not have a passport.

21. The probability that the tax-limit for income of males increases in the budget of a year is 0.66 and the probability that the tax-limit increases for income of females is 0.72. The probability that the tax-limit increases for income of both the males and females is 0.47. Find the probability that

(1) the tax-limit increases for income of only one of the two, males and females.

(2) the tax-limit does not increase for income of males as well as females in the budget of that year.

22. The price of petrol rises in 80% of the cases and the price of diesel rises in 77% of the cases after the rise in price of crude oil. The price of petrol and diesel rises in 68% cases. Find the probability that the price of diesel rises under the condition that there is a rise in the price of petrol.

23. As per the prediction of weather bureau, the probabilities for rains on three days, Thursday, Friday and Saturday in the next week are 0.8, 0.7 and 0.6 respectively. Find the probability that it rains on at least one of the three days in the next week.

(**Note :** The events of rains on three days, Thursday, Friday and Saturday of a week are independent.)

Section D

Answer the following questions :

1. 6 *LED* televisions and 4 *LCD* televisions are displayed in digital store A whereas 5 *LED* televisions and 3 *LCD* televisions are displayed in digital store B. One of the two stores is randomly selected and one television is selected from that store. Find the probability that it is an *LCD* television.

2. One number is randomly selected from the natural number 1 to 100. Find the probability that the number selected is either a single digit number or a perfect square.

3. A balanced coin is tossed thrice. If the first two tosses have resulted in tail, find the probability that tail appears on the coin in all the three trials.

4. If events A, B and C are independent events and $P(A)=P(B)=P(C)=p$ then find the value of $P(A \cup B \cup C)$ in terms of p.

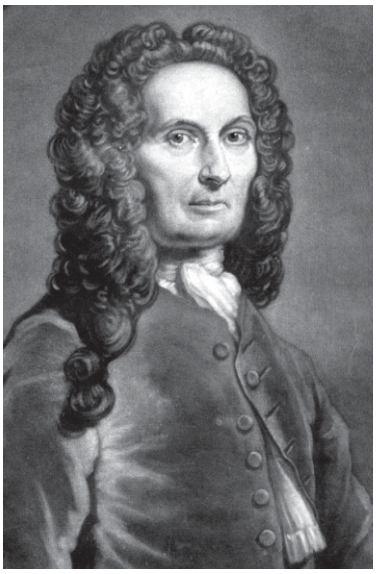
5. The genderwise data of a sample of 6000 employees selected from class 3 and class 4 employees in the government jobs of a state are shown in the following table :

Class of Employees	Gender		Total
	Males	Females	
Class 3	3600	900	4500
Class 4	400	1100	1500
Total	4000	2000	6000

One employee is randomly selected from all the class 3 and class 4 employees in government jobs of this state.

(1) If the selected employee is a male, find the probability that he belongs to class 3.

(2) If it is given that the selected employee belongs to class 3, find the probability that he is a male.



Abraham de Moivre
(1667 -1754)

Abraham de Moivre was a French mathematician known for de Moivre's formula, one of those that link complex numbers and trigonometry, and for his work on the normal distribution and probability theory. De Moivre wrote a book on probability theory, The Doctrine of Chances. De Moivre first discovered Binet's formula, the closed-form expression for Fibonacci numbers linking the n^{th} power of the golden ratio ϕ to the n^{th} Fibonacci number. He also was the first to postulate the Central Limit Theorem, a cornerstone of probability theory. In the later editions of his book, de Moivre included his unpublished result of 1733, which is the first statement of an approximation to the binomial distribution in terms of what we now call the normal or Gaussian function.

De Moivre continued studying the fields of probability and mathematics until his death and several additional papers were published after his death.

2

Random Variable and Discrete Probability Distribution

Contents :

2.1 Random Variable

We have studied about random experiment, sample space and probability in the chapter of probability. In this chapter, we shall study random variable and discrete probability distribution.

First of all, we shall define random variable and then we shall understand it by illustration.

Random Variable : Let U be a sample space of a random experiment. Every element of U need not always be a number. However, we wish to assign a specific number to each outcome.

A function associating a real number with each outcome of U is called a random variable. It is denoted by X. That is, a random variable based on a sample space U is denoted by $X : U \to R$.

For example,

(i) The number of heads (H) in tossing an unbiased coin three times

(ii) The number of accidents during a week in a city

(iii) The weight of a person (in kilogram)

(iv) The maximum temperature of a day at a particular place (in Celsius)

Now, let us understand the concept of random variable by some illustrations.

(1) A balanced die is tossed once. If the number observed on the die is denoted by *'u'* then the elements of the sample space U of this experiment can be shown in the notation of a set as follows :

$$U = \{u \mid u = 1, 2, 3, 4, 5, 6\}$$

That is $U = \{1, 2, 3, 4, 5, 6\}$

If we associate a real number X with element u of sample space by

$X(u)$ = the number obtained on the die then we can write

$$X(u) = u, \; u = 1, 2, 3, 4, 5, 6$$

Thus, variable X will be a random variable assuming values 1, 2, 3, 4, 5 and 6.

In the above illustration, the element of U are numeric. Now, we consider an illustration in which the elements of U are non-numeric.

(2) Suppose a box contains four balls : one red, one blue, one yellow and one white ball. We donote the red ball by R, the blue ball by B, the yellow ball by Y and the white ball by W. A person draws three balls at a time at random from the box. The sample space associated with this experiment is

$$U = \{RBY, RBW, BYW, WYR\}$$

Suppose for the element u of U,

$X(u) =$ the number of white ball in u then $X(RBY) = 0$, $X(RBW) = 1$, $X(BYW) = 1$, $X(WYR) = 1$.

Thus, random variable X assumes the values in the set $\{0, 1\}$. The outcomes of this sample space are not in numbers but we associate them with real numbers by a random variable.

(3) Suppose the heights of students in a class lie between 120 cm and 180 cm. If we measure the height of a student of this class then it will assume any value between 120 cm and 180 cm.

Here, the sample space is $U = \{u \mid 120 \le u \le 180\}$.

If we denote the height of a selected student by X then $X(u) = u =$ the height (in cm) of a selected student. Thus, X becomes a random variable which will be denoted as $X = x$, $120 \le x \le 180$.

In the above example (1) and example (2), random variable X assumes particular countable values whereas in example (3), random variable X can assume any value in the interval [120, 180]. This random variable differs from the random variables in earlier two examples.

Now, we shall understand the difference between these random variables in the following section.

2.1.1 Discrete Random Variable

A random variable X which can assume a finite or countable infinite number of values in the set R of real numbers is called a discrete random variable.

For example (i) Birth year of a randomly selected student.

(ii) Number of broken eggs in a box of 6 eggs.

Now, we shall understand about the discrete random variable by some specific examples.

(1) Suppose there is one black and two white balls in a box. Suppose the black ball is denoted by B and two white balls by W_1 and W_2. A person can play the following game by paying ₹ 15.

The person playing a game is asked to select two balls randomly with replacement from the box. He is paid an amount according to the colour of the balls selected by him as per the following conditions :

If a white ball is selected then ₹ 5 are paid for each selected white ball and if a black ball is selected then ₹ 15 are paid per black ball.

If we denote the net amount earned (amount received – amount paid for the game) by the player corresponding to each outcome of the experiment by X then X becomes a discrete random variable. The values assumed by the variable X are denoted in the following table :

Outcome of the experiment (Event)	The amount received by the person by playing the game	The amount paid to play the game	The value of X (in ₹)
W_1W_1	5 + 5 = 10	15	$X(W_1W_1)$ = 10 – 15 = –5
W_1W_2	5 + 5 = 10	15	$X(W_1W_2)$ = 10 – 15 = –5
W_1B_1	5 + 15 = 20	15	$X(W_1B_1)$ = 20 – 15 = 5
W_2W_1	5 + 5 = 10	15	$X(W_2W_1)$ = 10 – 15 = –5
W_2W_2	5 + 5 = 10	15	$X(W_2W_2)$ = 10 – 15 = –5
W_2B_1	5 + 15 = 20	15	$X(W_2B_1)$ = 20 – 15 = 5
B_1W_1	15 + 5 = 20	15	$X(B_1W_1)$ = 20 – 15 = 5
B_1W_2	15 + 5 = 20	15	$X(B_1W_2)$ = 20 – 15 = 5
B_1B_1	15 + 15 = 30	15	$X(B_1B_1)$ = 30 - 15 = 15

Thus, the random variable X assumes the values –5, 5 and 15 only. That is the total number of values of X is finite.

(2) Suppose a coin is tossed until either a tail (T) or four heads (H) occur. Let X denote the number of tosses required.

The sample space associated with this random experiment is

$U = \{T, HT, HHT, HHHT, HHHH\}$

The random variable X denotes the number of tosses required for the coin associated with the experiment and it assumes any one value out of 1, 2, 3 and 4 for the sample points of the sample space.

$$X(T) = 1,\ X(HT) = 2,\ X(HHT) = 3$$

$$X(HHHT) = 4,\ X(HHHH) = 4$$

The discrete random variable X assumes the finite number of values.

(3) Consider the random variable X denoting the number of tails before getting the first head in the experiment of tossing a coin till the first head is obtained.

In this experiment, head will appear either in the first trial or in the second trial or in the third trial and so on... Similarly, the first head may be obtained after tossing a coin infinite times. Hence, the sample space associated with random experiment becomes

$$U = \{H, TH, TTH, TTTH, TTTTH, ...\}$$

Thus, the number of tails before getting the first head will be 0, 1, 2, 3, 4....

Thus, the random variable X assumes any one value from the countable infinite number of values 0, 1, 2, 3, 4....

2.1.2 Continuous Random Variable

A random variable X which can assume any value in R, the set of real numbers or in any interval of R is called a continuous random variable.

For example (i) The actual amount of coffee in a coffee mug having a capacity of 250 millilitre.

(ii) Waiting time for a lift on any one floor of a high-rise office building.

Now, we shall understand more about the continuous random variable by the following examples.

(1) Denote the time taken by a student to finish a test of 3 hours duration by random variable X. The sample space here is

$$U = \{u | 0 \le u \le 3\}.$$

Since the time taken by any student for the exam takes any real value from 0 to 3 and the random variable X, the actual time taken by a student to complete the exam, will also be any real value from 0 to 3.

Thus,

$$X(u) = u, \ 0 \le u \le 3.$$

That means $X = x, 0 \le x \le 3$

The random variable X assumes any real value from 0 to 3, which is a subset of R and hence X is a continuous random variable.

(2) Suppose there are two stations A and B on an express highway. The distance of station B from station A is 200 km. Let us consider an experiment to know the place of accident between two stations A and B. For the sake of simplicity, let us fix the position of station A at 0 km and of station B at 200 km. The sample space of this experiment is any real value between 0 to 200. So, we can write the sample space for this experiment as

$$U = \{u \mid 0 \le u \le 200\}$$

Suppose the random variable X denotes the distance (in kilometer) of the place of the accident between two stations A and B from the station A. Then the random variable X is defined as below :

X (u) = distance of the place of accident from the station A.

In short, we can define the random variable X as $X = x, \ 0 \le x \le 200$.

The random variable X assumes any real value from 0 to 200, which is subset of R, the set of real numbers. So, X is a continuous random variable.

2.2 Discrete Probability Distribution

Suppose $X : U \rightarrow R$ is a random variable which assumes all the values of the subset $\{x_1, x_2, ..., x_n\}$ of *R*. Further, suppose *X* assumes a value x_i with probability $P(X = x_i) = p(x_i)$. If $p(x_i) > 0, \ i = 1, 2, ..., n$ and $\Sigma\, p(x_i) = 1$ then the set of real values $\{x_1, x_2, ..., x_n\}$ and $\{p(x_1), p(x_2), ..., p(x_n)\}$ is called the discrete probability distribution of a random variable *X*. The discrete probability distribution of a random variable *X* is expressed in a tabular form as follow :

$X = x$	x_1	x_2	....	x_i	...	x_n
$p(x)$	$p(x_1)$	$p(x_2)$	....	$p(x_i)$	...	$p(x_n)$

Here, $0 < p(x_i) < 1, i = 1, 2, ..., n$ and $\Sigma\, p(x_i) = 1$

2.2.1 Illustrations for Probability Distribution of Discrete Variable

Illustration 1 : Determine whether the values given below are appropriate as the values of a probability distribution of a discrete random variable *X*, which assumes the values 1, 2, 3 and 4 only.

(i) $p(1) = 0.25, p(2) = 0.75, p(3) = 0.25, p(4) = -0.25$

(ii) $p(1) = 0.15, p(2) = 0.27, p(3) = 0.29, p(4) = 0.29$

(iii) $p(1) = \frac{1}{19}, \quad p(2) = \frac{9}{19}, \quad p(3) = \frac{3}{19}, \quad p(4) = \frac{4}{19}$

(i) The value of $P(4)$ is –0.25, which is negative. It does not satisfy the condition $p(x_i) > 0, i = 1, 2, 3, 4$ of discrete probability distribution. So, given values are not suitable for the probability distribution of a discrete variable. Thus, the given distribution cannot be called a probability distribution of a discrete variable.

(ii) For every value 1, 2, 3 and 4 of *X*, $p(x) > 0$, and $p(1) + p(2) + p(3) + p(4) = 1$. Thus, both the conditions of probability distribution of discrete variable are satisfied. So, the given values are appropriate and the given distribution is probability distribution of a discrete variable.

(iii) Here $p(x_i) > 0, i = 1, 2, 3, 4$ but, sum of probabilities

i.e. $p(1) + p(2) + p(3) + p(4) = \frac{17}{19}$, is not 1. So, the given values are not appropriate for the probability distribution. So, the given distribution cannot be called a probability distribution of discrete variable.

Illustration 2 : Determine when the following distribution is a probability distribution of discrete variable. Hence obtain the probability for $x = 2$:

$$p(x) = c\left(\frac{1}{4}\right)^x, \quad x = 1, 2, 3, 4$$

Here, $p(1) = c\left(\frac{1}{4}\right), \ p(2) = c\left(\frac{1}{4}\right)^2 = c\left(\frac{1}{16}\right), \ p(3) = c\left(\frac{1}{4}\right)^3 = c\left(\frac{1}{64}\right), \ p(4) = c\left(\frac{1}{4}\right)^4 = c\left(\frac{1}{256}\right)$

Now, total probability should be 1 for a discrete probability distribution.

$p(1)+p(2)+p(3)+p(4)=1$

$\therefore\ c\left(\frac{1}{4}\right)+c\left(\frac{1}{16}\right)+c\left(\frac{1}{64}\right)+c\left(\frac{1}{256}\right)=1$

$\therefore\ c\left[\frac{1}{4}+\frac{1}{16}+\frac{1}{64}+\frac{1}{256}\right]=1$

$\therefore\ c\left[\frac{85}{256}\right]=1$

$\therefore\ c=\frac{256}{85}$

Thus, when $c=\frac{256}{85}$, the given distribution becomes probability distribution of a discrete variable.

Now, $P(X=2)=c\left(\frac{1}{4}\right)^2$

$=\frac{256}{85}\times\frac{1}{16}$

$=\frac{16}{85}$

$\therefore$ The probability of $X=2$ is $\frac{16}{85}$.

Illustration 3 : A random variable X denotes the number of accidents per year in a factory and the probability distribution of X is given below :

$X=x$	0	1	2	3	4
$p(x)$	$4K$	$15K$	$25K$	$5K$	K

(i) Find the constant K and rewrite the probability distribution.

(ii) Find the probability of the event that one or two accidents will occur in this factory during the year.

(iii) Find the probability that no accidents will take place during the year in the factory.

(i) By the definition of discrete probability distribution, we must have

$p(0)+p(1)+p(2)+p(3)+p(4)=1$

That is $4K+15K+25K+5K+K=1$

$\therefore 50K=1$

$\therefore K=\frac{1}{50}$

$=0.02$

Thus, when $K=0.02$, the given distribution becomes a probability distribution of a discrete variable, which is given below :

$X=x$	0	1	2	3	4	Total
$p(x)$	0.08	0.30	0.50	0.10	0.02	1

(ii) Probability of occurrence of one or two accidents

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

$= 0.30 + 0.50$

$= 0.80$

(iii) Probability that accidents do not occur :

$= P(X=0)$

$= 0.08$

Illustration 4 : In a factory, packets of produced blades are prepared having 50 blades in each packet. A quality control engineer randomly selects a packet from these packets and examines all the blades of the selected packet. If 4 or more defective blades are observed in the selected packet then the packet is rejected. The probability distribution of the defective blades in the packet is given below :

Number of defective blades in the packet	0	1	2	3	4	5	6 or more
Probability	$9K$	$3K$	$3K$	$2K$	$2K$	$K-0.02$	0.02

From the given probability distribution,

(i) Find constant *K*.

(ii) Find the probability that the randomly selected packet is accepted by the quality control engineer.

(i) Let X = number of defective blades found during the inspection of the packet.

By definition of discrete probability distribution

$p(0)+p(1)+p(2)+p(3)+p(4)+p(5)+p(6 \text{ or more})=1.$

$\therefore\ 9K + 3K + 3K + 2K + 2K + K - 0.02 + 0.02 = 1$

$\therefore\ 20K = 1$

$\therefore\ K = \frac{1}{20} = 0.05$

(ii) The randomly selected packet is accepted by the quality control engineer only when 3 or less defective blades are found in the packet.

$\therefore\ P(X \le 3)$

$= p(0)+p(1)+p(2)+p(3)$

$= 9K + 3K + 3K + 2K$

$= 17K$

$= 17(0.05)$

$= 0.85 \quad (\because K = 0.05)$

Illustration 5 : There are 4 red and 2 white balls in a box. 2 balls are drawn at random from the box without replacement. Obtain probability distribution of number of white balls in the selected balls.

Suppose X denotes the number of white balls in the selected two balls. X may assume the values 0, 1 and 2.

$X = 0$ means there will not be any white balls in the selected two balls that means both the selected balls are red.

$$\therefore P(X = 0) = P\,(2 \text{ red balls}) = \frac{{}^4C_2}{{}^6C_2} = \frac{6}{15}$$

Now, $x = 1$ means there will be one white ball and one red ball in the two selected balls.

$$\therefore P(X = 1) = P\,(1 \text{ White ball, } 1 \text{ Red ball})$$

$$= \frac{{}^2C_1 \times {}^4C_1}{{}^6C_2}$$

$$= \frac{2 \times 4}{15} = \frac{8}{15}$$

And $X = 2$ means both the selected ball will be white.

$$\therefore P(X = 2) = P\,(2 \text{ White balls})$$

$$= \frac{{}^2C_2}{{}^6C_2}$$

$$= \frac{1}{15}$$

Thus, probability distribution of random variable X can be written as follows :

$X = x$	0	1	2
$p(x)$	$\frac{6}{15}$	$\frac{8}{15}$	$\frac{1}{15}$

$p(x) > 0$ and $\Sigma p(x) = 1$

2.2.2 Mean and Variance

Now, we will discuss two important results based on the probability distribution of discrete random variable. One of them is expected value (mean) of the random variable and the other is variance of the random variable.

Let X be a discrete random variable which assumes one of the values $x_1, x_2, ..., x_n$ only and its probability distribution is as follows :

$X = x$	x_1	x_2	....	x_i	...	x_n
$p(x)$	$p(x_1)$	$p(x_2)$	....	$p(x_i)$	...	$p(x_n)$

Where $0 < p(x_i) < 1,\ i = 1, 2, ..., n$ and $\Sigma\, p(x_i) = 1$

The mean of discrete random variable is denoted by μ or $E(X)$. It is defined as follows :

$$\mu = E(X) = \Sigma\, x_i\, p(x_i)$$

This value is also called expected value of discrete variable X.

The variance of discrete random variable X is denoted by σ^2 or $V(X)$, which is defined as follows :

$$\sigma^2 = V(X) = E(X-\mu)^2$$

$$= E(X^2) - (\mu)^2$$

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

Where $E(X^2) = \Sigma\, x_i^2\, p(x_i)$

Note : (i) We will use the following notations for the sake of simplicity.

$\Sigma\, x\, p(x)$ instead of $\Sigma\, x_i\, p(x_i)$

and

$\Sigma\, x^2 p(x)$ instead of $\Sigma\, x_i^2\, p(x_i)$

(ii) The mean and variance of variable X are also called mean and variance of the distribution of X respectively.

(iii) The value of the variance of variable X is always positive.

We consider the following examples to find mean and variance of the discrete probability distribution.

Illustration 6 : Find constant C for the following discrete probability distribution. Hence obtain mean and variance of this distribution.

$$\boldsymbol{p(x) = C \cdot {}^4P_x,\ x = 0, 1, 2, 3, 4}$$

From the property of discrete probability distribution we must have

$$p(0) + p(1) + p(2) + p(3) + p(4) = 1$$

$$\therefore\ C \cdot {}^4P_0 + C \cdot {}^4P_1 + C \cdot {}^4P_2 + C \cdot {}^4P_3 + C \cdot {}^4P_4 = 1$$

$$\therefore C\left[\frac{4!}{4!} + \frac{4!}{3!} + \frac{4!}{2!} + \frac{4!}{1!} + \frac{4!}{0!}\right] = 1$$

$$\therefore C[1 + 4 + 12 + 24 + 24] = 1$$

$$\therefore C[65] = 1$$

$$\therefore C = \frac{1}{65}$$

Thus, the probability distribution can be written in the tabular form as follow :

$X = x$	0	1	2	3	4	Total
$p(x)$	$\frac{1}{65}$	$\frac{4}{65}$	$\frac{12}{65}$	$\frac{24}{65}$	$\frac{24}{65}$	1

Now, mean of the distribution $= \mu = \Sigma xp(x)$

$$= 0\left(\frac{1}{65}\right) + 1\left(\frac{4}{65}\right) + 2\left(\frac{12}{65}\right) + 3\left(\frac{24}{65}\right) + 4\left(\frac{24}{65}\right)$$

$$= \frac{0+4+24+72+96}{65}$$

$$= \frac{196}{65}$$

Now, we obtain $E\left(X^2\right)$.

$$E\left(X^2\right) = \Sigma\, x^2 p(x)$$

$$= (0)^2\left(\frac{1}{65}\right) + (1)^2\left(\frac{4}{65}\right) + (2)^2\left(\frac{12}{65}\right) + (3)^2\left(\frac{24}{65}\right) + (4)^2\left(\frac{24}{65}\right)$$

$$= 0 + \frac{4}{65} + \frac{48}{65} + \frac{216}{65} + \frac{384}{65}$$

$$= \frac{652}{65}$$

Hence, variance of the distribution $= V(X)$

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

$$= \frac{652}{65} - \left(\frac{196}{65}\right)^2$$

$$= \frac{42380 - 38416}{4225} = \frac{3964}{4225}$$

Illustration 7 : There are two red and one green balls in a box. Two balls are drawn at random with replacement from the box. Obtain probability distribution of number of red balls in the two balls drawn and find its mean and variance.

Let us denote the number of red balls in the selected two balls by X. Then we obtain the probability distribution of X as follow.

Let us denote the two red balls of the box by R_1 and R_2 and green ball by G.

The number of red balls in the selected balls and its probability can be obtained as in the following table.

Selected two balls (Event)	**Probability of the event**	$X = x$
R_1R_1	$\frac{1}{9}$	2
R_1R_2	$\frac{1}{9}$	2
R_1G	$\frac{1}{9}$	1
R_2R_1	$\frac{1}{9}$	2
R_2R_2	$\frac{1}{9}$	2
R_2G	$\frac{1}{9}$	1
GR_1	$\frac{1}{9}$	1
GR_2	$\frac{1}{9}$	1
GG	$\frac{1}{9}$	0

From the above table we can say that :

(i) Probability of getting 0 red ball

$= P(X = 0)$

$= \frac{1}{9}$

(ii) Probability of getting 1 red ball

$= P(X = 1)$

$= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9}$

$= \frac{4}{9}$

(iii) Probability of getting 2 red balls

$= P(X = 2)$

$= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9}$

$= \frac{4}{9}$

Thus, the probability distribution of X can be written in the tabular from as follow :

$X = x$	0	1	2	Total
$p(x)$	$\frac{1}{9}$	$\frac{4}{9}$	$\frac{4}{9}$	1

Now, mean of the distribution $= \mu = E(X)$

$$= \Sigma x\, p(x)$$

$$= 0\left(\frac{1}{9}\right) + 1\left(\frac{4}{9}\right) + 2\left(\frac{4}{9}\right)$$

$$= \frac{0 + 4 + 8}{9}$$

$$= \frac{12}{9}$$

Now, we first find $E\left(X^2\right)$ to obtain variance of the distribution.

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

$$= 0^2\left(\frac{1}{9}\right) + 1^2\left(\frac{4}{9}\right) + 2^2\left(\frac{4}{9}\right)$$

$$= \frac{0+4+16}{9}$$

$$= \frac{20}{9}$$

So, using the formula $V(X) = E\left(X^2\right) - \left(E(X)\right)^2$,

$$V(X) = \frac{20}{9} - \left(\frac{12}{9}\right)^2$$

$$= \frac{20}{9} - \frac{144}{81}$$

$$= \frac{180 - 144}{81}$$

$$= \frac{36}{81}$$

Illustration 8 : There are 2 black and 2 white balls in a box. Two balls are drawn without replacement from it. Obtain probability distribution of the number of white balls in the selected balls. Hence find its mean and variance.

Suppose $X =$ number of white balls in the selected two balls then by the formula of probability

(i) Probability of $X = 0$

$$= P(X = 0) = P \text{ (0 white balls)} = \frac{{}^2C_0}{{}^4C_2} = \frac{1}{6}$$

(ii) Probability of $X = 1$

$= P(X=1) = P$ (1 white ball and 1 black ball)

$= \dfrac{{}^2C_1 \times {}^2C_1}{{}^4C_2}$

$= \dfrac{2\times 2}{6}$

$= \dfrac{4}{6}$

(iii) Probability of $X = 2$

$= P(X=2) = P$ (2 white balls)

$= \dfrac{{}^2C_2}{{}^4C_2}$

$= \dfrac{1}{6}$

Thus, the probability distribution of random variable X can be written in the tabular form as,

$X = x$	0	1	2	Total
$p(x)$	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{1}{6}$	1

Now, mean of the probability distribution $= E(X)$

$= \Sigma\, x\, p(x)$

$= 0\left(\frac{1}{6}\right) + 1\left(\frac{4}{6}\right) + 2\left(\frac{1}{6}\right)$

$= \dfrac{0+4+2}{6}$

$= 1$

Now, to obtain variance of the probability distribution, we first find $E\left(X^2\right)$.

$E\left(X^2\right) = \Sigma\, x^2 p(x)$

$= 0^2\left(\frac{1}{6}\right) + 1^2\left(\frac{4}{6}\right) + 2^2\left(\frac{1}{6}\right)$

$= \dfrac{0+4+4}{6}$

$= \dfrac{8}{6}$

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

$= \frac{8}{6} - (1)^2 \qquad (\because E(X) = 1)$

$= \dfrac{8-6}{6}$

$= \dfrac{2}{6}$

$= \dfrac{1}{3}$

Illustration 9 : Let X denote the maximum integer among the outcomes of tossing two dice simultaneously. Obtain the probability distribution of variable X and find its mean and variance.

By tossing two dice simultaneously, we have 36 events in the sample space U and the maximum integer of the outcomes will be one of the numbers 1, 2, 3, 4, 5 or 6. The following table gives the possible outcomes for variable X and the corresponding probability :

Event u of U	**Maximum integer** $X(u) = x$	$P(X = x)$
(1, 1)	1	$\frac{1}{36}$
(1, 2), (2, 1), (2, 2)	2	$\frac{3}{36}$
(1, 3), (2, 3), (3, 3) (3, 2), (3, 1)	3	$\frac{5}{36}$
(1, 4), (2, 4), (3, 4), (4, 4) (4, 3), (4, 2), (4, 1)	4	$\frac{7}{36}$
(1, 5), (2, 5), (3, 5), (4, 5), (5, 5) (5, 4), (5, 3), (5, 2), (5, 1)	5	$\frac{9}{36}$
(1, 6), (2, 6), (3, 6), (4, 6), (5, 6) (6, 6), (6, 5), (6, 4), (6, 3), (6, 2) (6, 1)	6	$\frac{11}{36}$
		Total 1

Now, mean of $X = E(X)$

$$= \Sigma x\, p(x)$$

$$= 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right)$$

$$= \frac{161}{36}$$

Now, $E\left(X^2\right) = \Sigma x^2 p(x)$

$$= 1^2\left(\frac{1}{36}\right) + 2^2\left(\frac{3}{36}\right) + 3^2\left(\frac{5}{36}\right) + 4^2\left(\frac{7}{36}\right) + 5^2\left(\frac{9}{36}\right) + 6^2\left(\frac{11}{36}\right)$$

$$= \frac{791}{36}$$

Variance of $X = V(X)$

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

$= \frac{791}{36} - \left(\frac{161}{36}\right)^2$

$= \frac{791}{36} - \frac{25921}{1296}$

$= \frac{791 \times 36 - 25921}{1296}$

$= \frac{28476 - 25921}{1296}$

$= \frac{2555}{1296}$

Illustration 10 : It is observed from the life table that the probability that a 40 years old man will live one more year is 0.95. Life insurance company wishes to sell one year life insurance policy of Rs. 10,000 to such a man. What should be the minimum premium of the policy so that expected gain of the company would be positive ?

Let X be the company's gain and yearly premium of the policy be ₹ K, $K > 0$. Then gain of the company is $X = K$ if 40 year old man will live for one year and gain of the company is $X = K - 10{,}000$ if 40 year old man will die within a year.

Thus, the probability distribution of the gain of the company is as follow :

$X = x$	K	$K - 10000$
$p(x)$	0.95	0.05

Hence, expected gain of the company

$= E(X)$

$= \Sigma x\, p(x)$

$= K(0.95) + (K - 10000)(0.05)$

$= K(0.95) + K(0.05) - 500$

$= K(0.95 + 0.05) - 500$

$= K - 500$

Now, for positive expected gain, we must have

$K - 500 > 0$

$\therefore K > 500$

So, the company should fix the premium more than ₹ 500 so that the expected gain of the company be will positive.

EXERCISE 2.1

1. Examine whether the following distribution is a probability distribution of a discrete random variable X :

$$p(x) = \frac{x+2}{25}, \quad x = 1, 2, 3, 4, 5$$

2. If the following distribution is a probability distribution of variable X then find constant K.

$$p(x) = \frac{6-|x-7|}{K}, \quad x = 4, 5, 6, 7, 8, 9, 10$$

3. The probability distribution of a random variable X is defined as follows :

$$p(x) = \frac{K}{(x+1)!}, \quad x = 1, 2, 3; K = \text{constant}$$

Hence find (i) constant K (ii) $P(1 < X < 4)$

4. The probability distribution of a random variable X is as follows :

$X = x$	-2	-1	0	1	2
$p(x)$	$\frac{K}{3}$	$\frac{K}{3}$	$\frac{K}{3}$	$2K$	$4K^2$

Then (i) determine acceptable value of constant K. (ii) Find the variance of the distribution.

5. The probability distribution of a random variable X is $P(x)$. Variable X can assume the values $x_1 = -2, x_2 = -1, x_3 = 1$ and $x_4 = 2$ and if $4P(x_1) = 2P(x_2) = 3P(x_3) = 4P(x_4)$ then obtain mean and variance of this probability distribution.

6. A die is randomly tossed two times. Determine the probability distribution of the sum of the numbers appearing both the times on the die and obtain expected value of the sum.

7. A box contains 4 red and 2 blue balls. Three balls are simultaneously drawn at random. If X denotes the number of red balls in the selected balls, find the probability distribution of X and find the expected number of red balls in the selected balls.

8. A coin is tossed till either a head or 5 tails are obtained. If a random variable X denotes the necessary number of trials of tossing the coin then obtain probability distribution of the random variable X and calculate its mean and variance.

9. A shopkeeper has 6 tickets in a box. 2 tickets among them are worth a prize of ₹ 10 and the remaining tickets are worth a prize of ₹ 5. If a ticket is drawn at random from the box, find the expected value of the prize.

*

2.3 Binomial Probability Distribution

In the earlier sections we considered continuous and discrete random variable and probability distribution of a discrete random variable. Now, we shall study an important probability distribution of a discrete random variable.

In some random experiments, there are only two outcomes. We call such outcomes as success and failure. These outcomes are mutually exclusive. We call such experiments as dichotomous experiments. The illustrations of some of these situations are given in the table below :

Experiment		Possible outcomes	
		Success	Failure
(i)	To know the effect of advertisement given to increase the sale of produced units	sale increased	sale did not increase
(ii)	To find the error in a letter typed by a type-writer	Error observed	Error not observed
(iii)	To know the effect of a drug on blood pressure given to the patients of high blood pressure	Blood pressure decreased	Blood pressure did not decrease
(iv)	To inspect whether produced item is defective	Item is defective	Item is not defective

If we denote the success by S and failure by F for such types of dichotomous experiment and the probabilities of such outcomes by p and q respectively then

$P(S) = p$ and $P(F) = q,\ 0 < p < 1,\ 0 < q < 1,\ p + q = 1$

Since there are only two outcomes of such an experiment and both are mutually exclusive, we have $p + q = 1$ and hence $q = 1 - p$.

If it is possible to repeat such a dichotomous random experiment n times and each repetition is done under identical conditions then the probability of success p remains constant in each trial. We call such experiments as Bernoulli Trials. Its actual definition can be given as follows :

Bernoulli Trials : Suppose dichotomous random experiment has two outcomes, success (S) and failure (F). If this experiment is repeated n times under identical conditions and the probability $p(0 < p < 1)$ of getting a success at each trial is constant then such trials are called Bernoulli Trials.

Properties of Bernoulli Trials

(1) The probability of getting a success at each Bernoulli trial remains constant.

(2) Bernoulli trials are mutually independent. That means getting success or failure at any trial does not depend on getting success or failure at the previous trial.

(3) Success and failure are mutually exclusive and exhustive events. Therefore $q = 1 - p$.

Binomial Probability Distribution

Suppose X denotes the number of successes in a sequence of success *(S)* and failure (F) obtained in n Bernoulli trials, then X is called a binomial random variables and X assumes any value in the finite set $\{0, 1, 2, ..., n\}$. The probability distribution of the binomial random variable X is defined by the following formula :

$$P(X = x) = p(x) = {}^nC_x p^x q^{n-x}, \ x = 0, 1, 2, ..., n, 0 < p < 1, q = 1 - p$$

This probability distribution is called binomial Probability Distribution. We shall call such a distribution in short as binomial distribution.

If positive integer n and probability of success p are known here, the whole probability distribution that means the probability of each possible value of X can be determined. Hence, n and p are called parameters of the binomial distribution. We denote binomial distribution having parameters n and p as $b(n, p)$.

Note : If we repeat an experiment having such Bernoulli trials N times and $p(x)$ is the probability of getting x successes in the experiment then expected frequency of number of successes in N repetitions $= N \cdot p(x)$

2.3.1 Properties of Binomial Distribution

(1) Binomial distribution is a discrete distribution.

(2) Its parameters are n and p.

(3) The mean of the distribution is np which denotes average (expected) number of successes in n Bernoulli trials.

(4) The variance of the distribution is npq and its standard deviation is $\sqrt{npq}$.

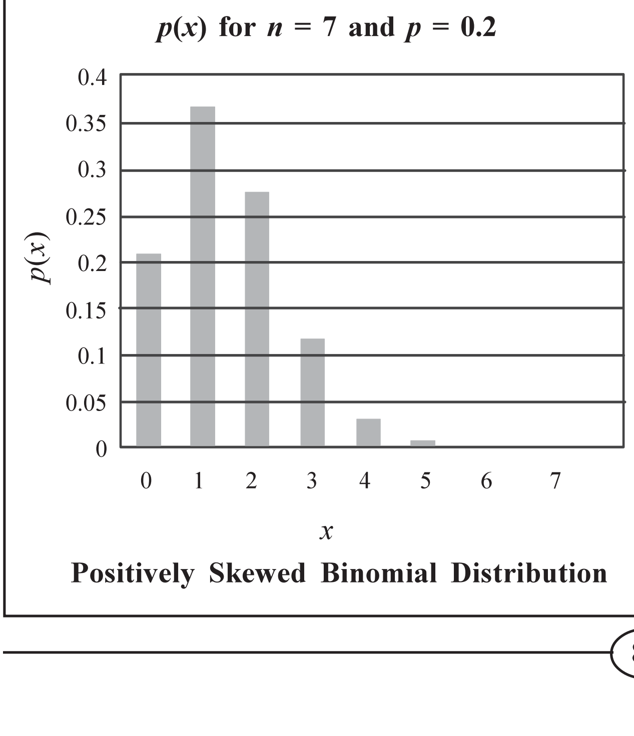
(5) For binomial distribution, mean is always greater than the variance and $\frac{\text{Variance}}{\text{Mean}} = q =$ probability of failure.

(6) If $p < \frac{1}{2}$ then the skewness of the distribution is positive for any value of n.

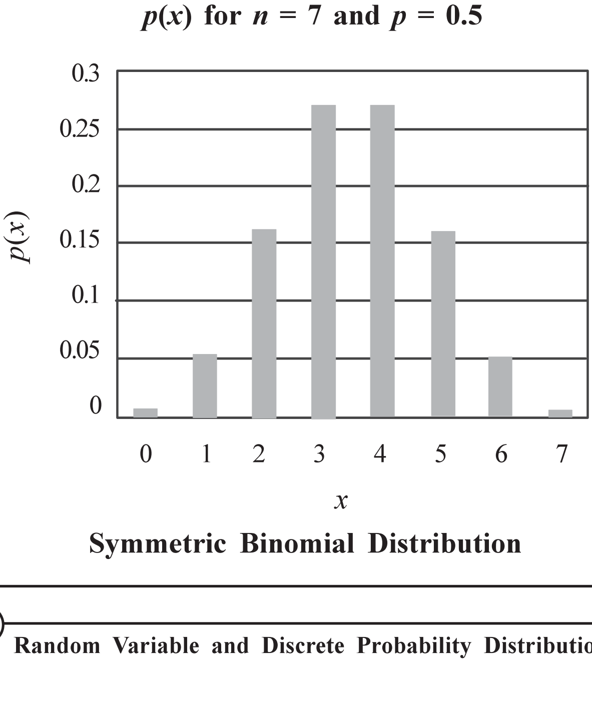
(7) If $p = \frac{1}{2}$ then the distribution becomes symmetric that means the skewness of the distribution is zero for any value of n.

(8) If $p > \frac{1}{2}$ then the skewness of the distribution is negative for any value of n.

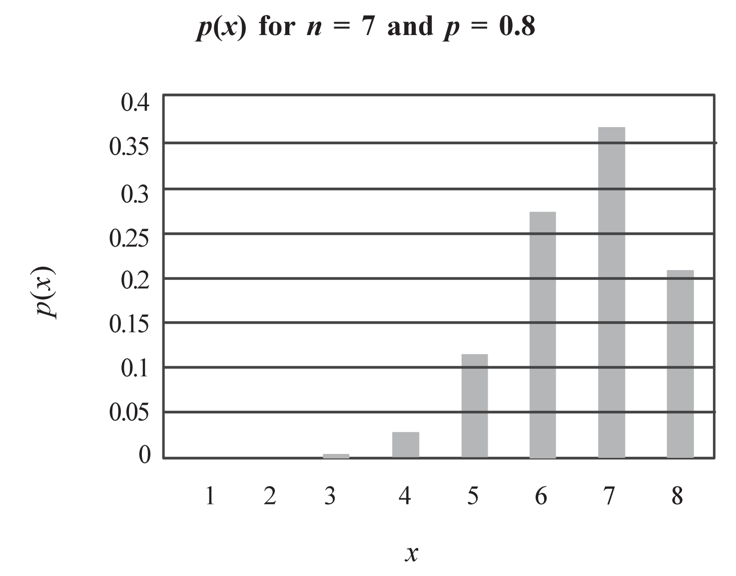
The properties (6), (7) and (8) can be clearly seen from the following graphs :



Positively Skewed Binomial Distribution



Symmetric Binomial Distribution



Negatively Skewed Binomial Distribution

2.3.2 Illustrations of Binomial Distribution

Illustration 11 : There are 3 % defective items in the items produced by a factory. 4 items are selected at random from the items produced. What is the probability that there will not be any detective item ?

If the event that the selected items is defective is considered as success then the probability of success $p = 0.03$ and $n = 4$. None of the selected items is defective means $X = 0$.

Now,

$$p(x) = {}^{n}c_{x} p^{x} q^{n-x}, \ x = 0, 1, 2, ..., n$$

Putting the values of *n, p,* $q = 1 - p$ and *x* in the formula,

$$P(X = 0) = {}^{4}c_{0} (0.03)^{0} (0.97)^{4-0}$$

$$= (0.97)^{4}$$

$$= 0.8853$$

Thus, the probability of getting no defective item in the selected 4 items is 0.8853.

Illustration 12 : The probability that a person living in a city is a non-vegetarian is 0.20. Find the probability of at the most two persons out of 6 persons randomly selected from the city is non-vegetarian.

If we consider the event that a person is non-vegetarian as success then we are given the probability of success $p = 0.20$ and $n = 6$.

If we take $X =$ number of non-vegetarians among the selected persons then the probability of $X \leq 2$

is obtained by putting the values of *n, p* and *x* in the formula of binomial probability distribution

$$p(x) = {}^{n}C_{x} p^{x} q^{n-x}, \ x = 0, 1, 2, ..., n$$

$$p(X \le 2) = p(X = 0 \ \text{ or } \ X = 1 \ \text{ or } \ X = 2)$$

$$= p(0) + p(1) + p(2)$$

$$= {}^{6}C_{0}(0.20)^{0}(0.80)^{6} + {}^{6}C_{1}(0.20)^{1}(0.80)^{6-1} + {}^{6}C_{2}(0.20)^{2}(0.80)^{6-2}$$

$$= 0.2621 + 6(0.20)(0.3277) + 15(0.04)(0.4096)$$

$$= 0.2621 + 0.3932 + 0.2458$$

$$= 0.9011$$

Illustration 13 : The mean and variance of a binomial distribution are 3.9 and 2.73 respectively. Find the number of Bernoulli trials conducted in this distribution and write $p(x)$.

Here, variance $= npq = 2.73$ and mean $= np = 3.9$.

$$\therefore q = \frac{\text{Variance}}{\text{Mean}} = \frac{2.73}{3.9} = 0.7 \ \text{ and } \ p = 1 - q = 0.3$$

$$\text{Now} \ n = \frac{np}{p} = \frac{\text{Mean}}{p} = \frac{3.9}{0.3} = 13$$

Thus, the number of Bernoulli trials conducted in this distribution is 13. Since $n = 13$, $p = 0.3$ and $q = 0.7$ in the distribution, its $p(x)$ can be written as follow :

$$p(x) = {}^{13}C_{x}(0.3)^{x}(0.7)^{13-x}, \ x = 0, 1, 2, ..., 13.$$

Illustration 14 : During a war, on an average one ship out of 9 got sunk in a certain voyage. Find the probability that exactly 5 out of a convoy of 6 ships would arrive safely.

Suppose X = the number of ships that arrive safely out of a convoy of 6 ships during a war.

$n =$ total number of ships in a convoy $= 6$

$p =$ probability that a ship arrives safely in a certain voyage $= \frac{8}{9}$

$\therefore$ The probability that exactly 5 out of a convoy of 6 ships would arrive safely can be obtained by putting corresponding values in the formula

$$p(x) = {}^{n}C_{x} p^{x} q^{n-x}, \ x = 0, 1, 2, ..., n,$$

$$p(5) = {}^{6}C_{5}\left(\frac{8}{9}\right)^{5}\left(\frac{1}{9}\right)^{1}$$

$$= 6\left(\frac{32,768}{59,049}\right)\left(\frac{1}{9}\right)$$

$$= \frac{196608}{531441}$$

$$= 0.3700$$

Illustration 15 : Assume that on an average one line out of 4 telephone lines remains busy between 2 pm and 3 pm on week days. Find the probability that out of 6 randomly selected telephone lines (i) not more than 3 (ii) at least three of them will be busy.

Suppose $p =$ the probability of the event that the selected telephone line remains busy between 2 pm to 3 pm = $\frac{1}{4}$

and $X =$ the number of busy telephone lines out of 6 telephone lines between 2 pm to 3 pm.

It is given here that $n = 6$.

(i) The event that not more than 3 lines out of 6 randomlly selected telephone lines will be busy is the event that 3 or less telephone lines will be busy.

That is $X \leq 3$.

$\therefore$ To find probability of this event we use the formula of binomial probability distribution

$$p(x) = {}^{n}C_{x} p^{x} q^{n-x},\ x = 0, 1, 2, ..., n$$

$\therefore \quad P(X \leq 3)$

$= P\ (X = 0 \text{ or } 1 \text{ or } 2 \text{ or } 3)$

$= 1 - P\ (X = 4 \text{ or } 5 \text{ or } 6)$

$= 1 - \left[p\ (4) + p(5) + p(6)\right]$

$$= 1 - \left[{}^{6}C_{4} \left(\tfrac{1}{4}\right)^{4} \left(\tfrac{3}{4}\right)^{2} + {}^{6}C_{5} \left(\tfrac{1}{4}\right)^{5} \left(\tfrac{3}{4}\right)^{1} + {}^{6}C_{6} \left(\tfrac{1}{4}\right)^{6} \left(\tfrac{3}{4}\right)^{0}\right]$$

$$= 1 - \left[15 \left(\tfrac{1}{256}\right)\left(\tfrac{9}{16}\right) + 6\left(\tfrac{1}{1024}\right)\left(\tfrac{3}{4}\right) + \left(\tfrac{1}{4096}\right)\right]$$

$$= 1 - \left[\tfrac{135}{4096} + \tfrac{18}{4096} + \tfrac{1}{4096}\right]$$

$$= 1 - \tfrac{154}{4096} = \tfrac{3942}{4096} = 0.9624$$

(ii) The probability that at least 3 telephone lines will be busy

$= P(X \geq 3)$

$= P\ (X = 3 \text{ or } 4 \text{ or } 5 \text{ or } 6)$

$= p(3) + p\ (4) + p(5) + p(6)$

Now, from the above calculations we will get the values of $p\ (4)$, $p(5)$ and $p(6)$. So, we first find the value of $p(3)$.

$$p(3) = {}^6C_3 \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^3 = 20\left(\frac{1}{64}\right)\left(\frac{27}{64}\right)$$

$$= \frac{540}{4096}$$

Now, from the values of $p(4)$, $p(5)$ and $p(6)$ in question (i) and the value of $p(3)$ we obtained,

$$P(X \geq 3) = \frac{540}{4096} + \frac{135}{4096} + \frac{18}{4096} + \frac{1}{4096}$$

$$= \frac{694}{4096} = 0.1694$$

Illustration 16 : The parameters of binomial distribution of a random variable X are $n = 4$ and $p = \frac{1}{3}$. State the probability distribution of X in a tabular form and hence find the value of $P(X \leq 2)$.

Here, parameters are $n = 4$ and $p = \frac{1}{3}$ $\therefore q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$

Substituting the values of the parameters in the formula of binomial distribution,

$p(x) = {}^nC_x \, p^x \, q^{n-x}$, $x = 0, 1, 2, ..., n$ we have $p(x) = {}^4C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{4-x}$, $x = 0, 1, 2, 3, 4$.

Now, we calculate the values of $p(x)$ by putting the different values of x as 0, 1, 2, 3 and 4.

$$p(0) = {}^4C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^{4-0} = \frac{16}{81}$$

$$p(1) = {}^4C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^{4-1} = 4\left(\frac{1}{3}\right)\left(\frac{8}{27}\right) = \frac{32}{81}$$

$$p(2) = {}^4C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{4-2} = 6\left(\frac{1}{9}\right)\left(\frac{4}{9}\right) = \frac{24}{81}$$

$$p(3) = {}^4C_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^{4-3} = 4\left(\frac{1}{27}\right)\left(\frac{2}{3}\right) = \frac{8}{81}$$

$$p(4) = {}^4C_4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^{4-4} = 1\left(\frac{1}{81}\right)1 = \frac{1}{81}$$

These can be put in the tabular form as follows :

$X = x$	0	1	2	3	4	Total
$p(x)$	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$	1

Now, $P(X \le 2)$

$= p(X=0)+p(X=1)+p(X=2)$

$= \frac{16}{81}+\frac{32}{81}+\frac{24}{81}$

$=\frac{72}{81}$

$=\frac{8}{9}$

Illustration 17 : In a binomial distribution, for $P(X=x)=p(x)$, $n=8$ and $2p(4)=5p(3)$. Find the probability of getting success in all the trials for this distribution.

Here, we have $2p(4)=5p(3)$ and $n=8$

$\therefore$ Putting $n=8$ in the formula of binomial distribution, we get,

$p(x)={}^8C_x p^x q^{8-x}$, $x = 0, 1, 2, ..., 8$

Putting the values of $p(4)$ and $p(3)$ from this formula in the given condition

$2p(4)=3p(3)$

$2 \times {}^8C_4\, p^4\, q^{8-4} = 5 \times {}^8C_3 p^3 q^{8-3}$

$\therefore\ 2 \times (70)\, p^4\, q^4 = 5 \times (56)\, p^3 q^5$

$\therefore\ 140\, p^4\, q^4 = 280 p^3 q^5$

$\therefore\ p=2q$

$\therefore\ p=2(1-p)$

$\therefore\ p=2-2p$

$\therefore 3p=2$

$\therefore\ p=\frac{2}{3}$ and $q=1-p=1-\frac{2}{3}=\frac{1}{3}$.

Now, getting success in all the trials means the event of getting 8 successes since we have total 8 trials..

The probability of this events is $p(8)$.

$\therefore\ p(8)={}^8C_8\left(\frac{2}{3}\right)^8\left(\frac{1}{3}\right)^{8-8}$

$=1\times\left(\frac{2}{3}\right)^8\times 1$

$=\frac{256}{6561}$

Thus, the probability of getting success in all the trials is $\frac{256}{6561}$.

Illustration 18 : For a binomial distribution, mean = 18 and variance = 4.5. Determine whether the skewness of this distribution is positive or negative.

Here, mean $= np = 18$ and variance $= npq = 4.5$

$\therefore q = \frac{\text{Variance}}{\text{Mean}} = \frac{4.5}{18} = 0.25 = \frac{1}{4}$

$\therefore p = 1 - \frac{1}{4} = \frac{3}{4}$

Since the value of p is greater than $\frac{1}{2}$, the skewness of binomial distribution will be negative.

Illustration 19 : A balacned die is tossed 7 times. If the event of getting a number 5 or more is called success and *X* denotes the number of success in 7 trials then (i) Write the probability distribution of *X*. (ii) Find the probability of getting 4 successes. (iii) Find the probability of getting at the most 6 successes.

The sample space associated with tossing of a balanced die once is $U = \{1, 2, 3, 4, 5, 6\}$ and probability of getting each number is $\frac{1}{6}$.

If the event of getting a number 5 or more is called success then probability of success

$p =$ probability of getting 5 or 6 on the die

$= \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

$\therefore q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$

Here, total number of trials is 7. $\therefore n = 7$

(i) Using the probability distribution of X $p(x) = {}^nC_x p^x q^{n-x}$, $x = 0, 1, 2, ..., n$,

$$p(x) = {}^7C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{7-x}, \quad x = 0, 1, 2, 3, 4, 5, 6, 7$$

(ii) Probability of getting 4 successes

$$p(4) = {}^7C_4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^{7-4}$$

$$= 35\left(\frac{1}{81}\right)\left(\frac{8}{27}\right)$$

$$= \frac{280}{2187}$$

(iii) Probability of getting at the most 6 successes

$= p(X \le 6)$

$= 1 - p(X > 6)$

$= 1 - p(X = 7) \quad \because x = 0, 1, 2, ..., 7$

$= 1 - {}^7C_7 \left(\frac{1}{3}\right)^7 \left(\frac{2}{3}\right)^{7-7}$

$= 1 - \frac{1}{2187} = \frac{2186}{2187}$

Illustration 20 : A social worker claims that 10 % of the young children in a city have vision problem. A sample survey agency takes a random sample of 10 young children from the city to test the claim. If at the most one young child is affected by the vision problem, the claim of the social worker is rejected. Find (i) the probability that the claim of the social worker is rejected (ii) the expected number of young children having vision problem in the randomly selected 10 young children.

Suppose p = probability that a young child has eye problem

= 0.10 (by accepting the claim of social worker)

And X = the number of young childrens having eye problem in the randomly selected 10 young children

Here, putting $n = 10$ and $p = 0.10$ in the formula

$p(x) = {}^nC_x \; p^x q^{n-x}, \quad x = 0, 1, 2, ..., n$

of binomial distribution,

$p(x) = {}^{10}C_x (0.10)^x (0.90)^{10-x}, \; x = 0, 1, 2, ..., 10$

(i) Probability that at the most one young child has eye problem

$= p(0) + p(1)$

$= {}^{10}C_0 (0.10)^0 (0.90)^{10-0} + {}^{10}C_1 (0.10)^1 (0.90)^{10-1}$

$= 0.3487 + 10 (0.10)(0.3874)$

$= 0.3487 + 0.3874$

$= 0.7361$

Now, sample survey agency rejects the claim of the social worker if at the most one young child has eye problem.

$\therefore$ Probability of rejecting the claim of social worker by the sample survey agency $= 0.7361$.

(ii) The expected number of young children having eye problem in the randomly selected 10 young child

$= E(X) = np$

$= 10 \times$ probability that the selected young child has eye problem

$= 10 \times 0.10$

$= 1$

Illustration 21 : An experiment is conducted to toss five balanced coins simultaneously. If we consider occurrence of head (H) on the coin as success then obtain probability distribution of the number of successes. If such an experiment is repeated 3200 times then obtain expected frequency distribution of the number of successes. For this distribution, obtain expected value of the number successes and also obtain its standard deviation.

Since the coins are balanced, probability of getting head will be $\frac{1}{2}$.

$p =$ probability of success

$=$ probability of getting head $= \frac{1}{2}$.

$\therefore q = 1 - p = \frac{1}{2}$

Here, $n =$ number of coins $= 5$, $x =$ the number of successes in tossing of five coins.

Putting the values of n, p and q in the formula

$p(x) = {}^{n}C_{x}\, p^{x} q^{n-x}, \; x = 0, 1, 2, ..., n$

of the binomial distribution

$$p(x) = {}^{5}C_{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{5-x}, \quad x = 0, 1, 2, ..., 5$$

$$= {}^{5}C_{x} \left(\frac{1}{2}\right)^{x+5-x}$$

$$= {}^{5}C_{x} \left(\frac{1}{2}\right)^{5}$$

$$= \frac{{}^{5}C_{x}}{32}, \quad x = 0, 1, 2, ..., 5$$

Now, using the above formula, we calculate the probability for each x and the frequency for the number of successes in 3200 repetition of the experiment = $3200 \times p(x)$, $x = 0, 1, 2, ...,5$.

We present the calculations in the following table :

x	$p(x)$	**Expected Frequency** $= \mathbf{N} \times p(x)$
0	$\frac{{}^5C_0}{32} = \frac{1}{32}$	$3200 \times \frac{1}{32} = 100$
1	$\frac{{}^5C_1}{32} = \frac{5}{32}$	$3200 \times \frac{5}{32} = 500$
2	$\frac{{}^5C_2}{32} = \frac{10}{32}$	$3200 \times \frac{10}{32} = 1000$
3	$\frac{{}^5C_3}{32} = \frac{10}{32}$	$3200 \times \frac{10}{32} = 1000$
4	$\frac{{}^5C_4}{32} = \frac{5}{32}$	$3200 \times \frac{5}{32} = 500$
5	$\frac{{}^5C_5}{32} = \frac{1}{32}$	$3200 \times \frac{1}{32} = 100$

(ii) Expected value of the number of successes

$= np$

$= 5\left(\frac{1}{2}\right) = 2.5$

(iii) Standard deviation of the number of successes

$= \sqrt{npq}$

$= \sqrt{5 \times \left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right)} = \sqrt{\frac{5}{4}} = \sqrt{1.25}$

$= 1.118$

Illustration 22 : An advertisement company claims that 4 out of 5 house wives do not identify the difference between two different brands of butter. To check the claim, 5000 house wives are divided in groups, each group of 5 house wives. If the claim is true, in how many groups among these groups (i) at the most one house wife (ii) only two house wives can identify the difference between two different brands of butter ?

As per the claim made by an advertising company, 4 out of 5 house wives do not identify the difference between two different brands of butter.

$\therefore$ Its probability is $\frac{4}{5}$

That means the probability of identifying the difference between two different brands of butter by a house wife $= \frac{1}{5}$.

Let p = probability that the selected house wife can identify the difference between two different brands of butter $= \frac{1}{5}$.

To test the claim, selected 5000 housewives are divided in groups randomly, with each group having 5 house wives. So, there will be 1000 such groups.

If we take $X =$ the number of house wives who identify the difference between two different brands of butter in a group, $x = 0, 1, ..., 5$.

Thus, we have $n = 5$, $p = \frac{1}{5}$, $q = \frac{4}{5}$.

Putting the above values in the formula of binomial distribution

$$p(x) = {}^{n}C_{x}\ p^{x}\, q^{n-x},\ x = 0, 1, 2, ..., n$$

We get the following $p(x)$

$$p(x) = {}^{5}C_{x} \left(\frac{1}{5}\right)^{x} \left(\frac{4}{5}\right)^{5-x}, \quad x = 0, 1, 2, ..., 5$$

Using this formula, we calculate the probabilities for different values of x and multiplying such probabilities by 1000 we get the number of groups out of 1000 groups in which 0, 1, 2, 3, 4 or 5 house wives can identify the difference between two different brands of the butter..

(i) The number of groups out of 1000 groups in which at the most one house wife can identify the difference between two different brands of butter

$$= 1000 \times \left[p(0) + p(1) \right]$$

$$= 1000 \times \left[{}^{5}C_{0} \left(\frac{1}{5}\right)^{0} \left(\frac{4}{5}\right)^{5-0} + {}^{5}C_{1} \left(\frac{1}{5}\right)^{1} \left(\frac{4}{5}\right)^{5-1} \right]$$

$$= 1000 \times \left[\frac{1024}{3125} + 5 \times \left(\frac{1}{5}\right) \times \left(\frac{256}{625}\right) \right]$$

$$= 1000 \times \left[\frac{1024}{3125} + \frac{256}{625} \right]$$

$= 1000 \times [0.32768 + 0.4096]$

$= 1000 \times [0.73728]$

$= 737.28$

$\simeq 737$ groups

(ii) The number of groups out of 1000 groups in which only two house wives can identify the difference between the two brands of butter.

$= 1000 \times p(2)$

$= 1000 \times {}^5C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{5-2}$

$= 1000 \times 10 \times \left(\frac{1}{25}\right) \times \left(\frac{64}{125}\right)$

$= 1000 \times \frac{640}{3125}$

$= 1000 \times 0.2048$

$= 204.8$

$\simeq 205$ groups

EXERCISE 2.2

1. For a symmetrical binomial distribution with $n = 8$, find $p(X \leq 1)$.

2. Mean of a binomial distribution is 5 and its variance is equal to the probability of success. Find the parameters of this distribution and hence find the probability of the event of getting none of the failures for this distribution.

3. A person has kept 4 cars to run on rent. The probability that any car is rented during the day is 0.6. Find the probability that more than one but less than 4 cars are rented during a day.

4. There are 200 farms in a Taluka. Among the bore wells made in these 200 farms of the Taluka, salted water is found in 20 farms. Find the probability of the event of not getting salted water in 3 out of 5 randomly selected farms from the Taluka.

5. An example is given to 6 students to solve. The probability of getting correct solution of the problem by any student is 0.6. Students are trying to solve the problem independently. Find the probability of getting the correct solution by only 2 out of the 6 students.

*

Summary

- **Random Variable :** A function associating a real number with each outcome of the sample space of a random experiment is called random variable.
- **Discrete Random Variable :** A random variable X which can assume a finite or countable infinite number of values in the set R of real numbers is called a discrete random variable.
- **Continuous Random Variable :** A random variable X which can assume any value in R, the set of real numbers or in any interval of R is called continuous random variable.
- **Discrete Probability Distribution :** Suppose $X : U \to R$ is a random variable which assumes all values of a finite set $\{x_1, x_2, ..., x_n\}$ of R. Also suppose X assumes a value x_i with probability $p(x_i)$. If $p(x_i) > 0$ for $i = 1, 2, ..., n$ and $\Sigma\, p(x_i) = 1$ then the set of real values $\{x_1, x_2, ..., x_n\}$ and $\{p(x_1), p(x_2), ..., p(x_n)\}$ is called the discrete probability distribution of a random variable X which is expressed in a tabular form as follows :

$X = x$	x_1	x_2	....	x_i	...	x_n	Total
$p(x)$	$p(x_1)$	$p(x_2)$	....	$p(x_i)$	...	$p(x_n)$	1

Here $0 < p(x_i) < 1,\ i = 1, 2, ..., n$

- **Bernoulli Trials :** Suppose dichotomous random experiment has two outcomes success (S) and failure (F). If this experiment is repeated under identical conditions and the probability $p(0 < p < 1)$ of getting success at each trial is constant then such trials are called Bernoulli Trials.
- **Binomial Random Variable :** Suppose X denotes the number of successes in the sequence of success (S) and failure (F) obtained in n Bernoulli trials then X is called a binomial random variable.
- **Binomial Probability Distribution :** The probability distribution of a binomial random variable X is called binomial probability distribution.

List of Formulae

(1) Mean of discrete probability distribution $= \mu$

$= E(X)$

$= \Sigma\, x\, p(x)$

(2) Variance of discrete probability distribution $= \sigma^2$

$= V(X)$

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

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

(3) Binomial Probability Distribution

$P(X = x) = p(x) = {}^nC_x\, p^x q^{n-x}$, $x = 0, 1, 2, \ldots, n.$

$0 < p < 1,\ q = 1 - p$

(4) Mean of binomial probability distribution $= np$

(5) Variance of binomial probability distribution $= npq$

(6) Standard deviation of binomial probability distribution $= \sqrt{npq}$

(7) If an experiment having Bernoulli trials repeats N times and $p(x)$ is the probability of getting x successes in the experiment then expected frequency of the number of successes in N repetitions $= N \cdot p(x)$

EXERCISE 2

Section A

Find the correct option for the following multiple choice questions :

1. Which variable of the following will be an illustration of discrete variable ?

(a) Height of a student (b) Weight of a student

(c) Blood Pressure of a student (d) Birth year of a student

2. Which variable of the following will be an illustration of continuous variable ?

(a) Number of accidents occuring at any place

(b) Number of rainy days during a year

(c) Maximum temperature during a day

(d) Number of children in a family

3. A random variable X assume the values –1, 0 and 1 with respective probability $\frac{1}{5}$, K and $\frac{1}{3}$, where $0 < K < 1$ and X does not assume any value other than these values. What will be the value of $E(X)$?

(a) $\frac{2}{5}$ (b) $\frac{3}{5}$ (c) $\frac{2}{15}$ (d) $\frac{3}{15}$

4. A random variable X assumes the values –2, 0 and 2 only with respective probabilities $\frac{1}{5}, \frac{3}{5}$ and K. If $0 < K < 1$, what will be the value of K ?

(a) $\frac{1}{5}$ (b) $\frac{4}{5}$ (c) $\frac{2}{5}$ (d) $\frac{3}{5}$

5. Mean and variance of a discrete probability distribution are 3 and 7 respectively. What will be $E\left(X^2\right)$ for this distribution ?

(a) 10 (b) 4 (c) 40 (d) 16

6. For the probability distribution of a discrete random variable, $E(X)=5$ and $E\left(X^2\right)=35$. What will be the variance of this distribution ?

(a) 40 (b) 30 (c) 20 (d) 10

7. For a positively skewed binomial distribution with n = 10, which of the following values might be the value of mean ?

(a) 5 (b) 3 (c) 9 (d) 7

8. For which value of x, the value of $p(x)$ of binomial distribution with parameters $n = 4$ and $p = \frac{1}{2}$ becomes maximum ?

(a) 0 (b) 2 (c) 3 (d) 4

9. The binomial distribution has mean 5 and variance $\frac{10}{7}$. What will be the type of this distribution ?

(a) Positively skewed (b) Negatively skewed

(c) Symmetric (d) Nothing can be said about the distribution

10. Which of the following is the formula of probability of an event of not getting a success in the binomial distribution with parameters n and p ?

(a) ${}^nC_0 p^n q^0$ (b) ${}^nC_0 p^0 q^n$ (c) ${}^nC_0 p q^n$ (d) ${}^nC_0 p^n q$

Section B

Answer the following questions in one sentence :

1. Define discrete random variable.
2. Define continuous random variable.
3. Define discrete probability distribution.
4. State the formula to find mean of discrete variable.
5. State the formula to find variance of discrete variable.
6. Mean of a symmetrical binomial distribution is 7. Find the value of its parameter n.
7. The parameters of a binomial distribution are 10 and $\frac{2}{5}$. Calculate its variance.
8. State the relation between the probability of success and failure in Bernoulli trials.
9. State the relation between mean and variance of binomial distribution.
10. The probability of failure in a binomial distribution is 0.6 and the number of trials in it is 5. Find the probability of success.

Section C

Answer the following questions :

1. The probability distribution of a random variable X is as follows :

X	2	3	4	5
$p(x)$	0.2	0.3	$4C$	C

Determine the value of constant C.

2. Calculate mean of the discrete probability distribution $p(x) = \begin{cases} \frac{x-1}{6}; & x = 2, 3 \\ \frac{1}{2}; & x = 4 \end{cases}$

3. The probability distribution of a random variable is as follows :

$p(x) = \frac{x+3}{10}, \quad x = -2, 1, 2$

Hence calculate $E(X^2)$.

4. If $n=4$ for a symmetrical binomial distribution then find $p(4)$.

5. Define Bernoulli trials.

6. For a binomial distribution, if probability of success is double the probability of failure and $n=4$ then find variance of the distribution.

7. Find the standard deviation of the binomial distribution having $n=8$ and probability of failure $\frac{2}{3}$.

8. Find parameters of the binomial distribution where mean = 4 and variance = 2.

9. For a binomial distribution with $n=10$ and $q-p=0.6$, find mean of this distribution.

10. For a binomial distribution, standard deviation is 0.8 and probability of failure is $\frac{2}{3}$, find the mean of this distribution.

Section D

Answer the following questions :

1. The probability distribution of a random variable X is as follows :

$$p(x)=\begin{cases} K(x-1); & x=2, 3 \\ K; & x=4 \\ K(6-x) \ ; & x=5 \end{cases}$$

Find the value of constant K and the probability of the event that variable X assumes even numbers.

2. The probability distribution of a random variable X is as follows :

$$p(x)=C(x^2+x),\ x=-2, 1, 2$$

Find the value of C and show that $p(2)=3p(-2)$.

3. The distribution of a random variable X is $p(x)=K\cdot {}^5P_x,\ x = 0, 1, 2, 3, 4, 5$

Find constant K and mean of this distribution.

4. What is discrete probability distribution ? State its properties.

5. State properties of binomial distribution.

6. In a game of hitting a target, the probability that Ramesh will fail in hitting the target is $\frac{2}{5}$. If he is given 3 trials to hit the target, find the probability of the event he hits the target successfully in 2 trials. State mean of this distribution.

7. A person is asked to select a number from positive integers 1 to 7. If the number selected by him is odd then he is entitled to get the prize. If he is asked to take 5 trials then find the probability of the event that he will be entitled to get a prize in only one trial.

8. The mean and variance of the binomial distribution are 2 and $\frac{6}{5}$ respectively. Find $p(1)$ and $p(2)$ for this binomial distribution.

9. 10 % apples are rotten in a box of apples. Find the probability that half of the 6 apples selected from the box with replacement will be rotten and find the variance of the number of rotten apples.

Section E

Solve the following :

1. The probability distribution of the monthly demand of laptop in a store is as follows :

Demand of laptop	1	2	3	4	5	6
Probability	0.10	0.15	0.20	0.25	0.18	0.12

Determine the expected monthly demand of laptop and find variance of the demand.

2. Two dice are thrown simultaneously once. Obtain the discrete probability distribution of the number of dice for which the number '6' comes up.

3. If the probability that any 50 year old person will die within a year is 0.01, find the probability that out of a group of 5 such persons

(i) none of them will die within a year

(ii) at least one of them will die within a year.

4. The probability that a student studying in 12th standard of science stream will get admission to engineering branch is 0.3. 5 students are selected from the students who studied in this stream. Find the probability of the event that the number of students admitted to engineering branch is more than the number of students who did not get admission to the engineering branch.

5. The probability that a bomb dropped from a plane over a bridge will hit the bridge is $\frac{1}{5}$. Two bombs are enough to destroy the bridge. If 6 bombs are dropped on the bridge, find the probability that the bridge will be destroyed.

6. Normally, 40 % students fail in one examination. Find the probability that at least 4 students in a group of 6 students pass in this examination.

7. There are 3 red and 4 white balls in a box. Four balls are selected at random with replacement from the box. Find the probability of the event of getting (i) 2 red balls and 2 white balls (ii) all four white balls among the selected balls using binomial distribution.

Section F

Solve the following :

1. There are one dozen mangoes in a box of which 3 mangoes are rotten. 3 mangoes are randomly selected from the box without replacement. If X denotes the number of rotten mangoes in the selected mangoes, obtain the probability distribution of X and hence find expected value and variance of the rotten mangoes in the selected mangoes.

2. It is known that 50 % of the students studying in the 10th standard have a habit of eating chocolate. To examine the information, 1024 investigators are appointed. Every investigator randomly selects 10 students from the population of such students and examines them for the habit of eating chocolate. Find the expected number of investigators who inform that less than 30 percent of the students have a habit of eating chocolate.

James Bernoulli
(1654 –1705)

James (Jacob) Bernoulli was born in Basel, Switzerland. He was one of the many prominent mathematicians in the Bernoulli family. Following his father's wish, he studied theology (divinity) and entered the ministry. But contrary to the desires of his parents, he also studied mathematics and astronomy. He travelled throughout Europe from 1676 to 1682; learning about the latest discoveries in mathematics and the sciences under leading figures of the time. He was an early proponent of Leibnizian calculus and had sided with Leibniz during the Leibniz-Newton calculus controversy. He is known for his numerous contributions to calculus, and along with his brother Johann, was one of the founders of the calculus of variations. However, his most important contribution was in the field of probability, where he derived the first version of the law of large numbers. He was appointed as professor of mathematics at the University of Basel in 1687, remained in this position for the rest of his life.

●

"Normal Distribution is father of all probability distributions. For larger sample size almost all theoritical distributions follow normal distribution".

– Unknown

3

Normal Distribution

Contents :

3.1 Normal Distribution : Introduction, Probability Density Function

In the previous chapter, we have studied the probability distribution for a discrete random variable. Now, we shall study the probability distribution for a continuous random variable. We know that, if a random variable *X* can assume any value of real set *R* or within any interval of real set *R* then it is called continuous random variable. If a random variable can assume any value between the definite interval *a* to *b* then it is denoted by $a < x < b$. A function for obtaining probability that a continuous random variable assumes value between specified interval is called probability density function of that variable and it satisfies the following two conditions:

(1) The probability that the value of random variable lies within the specified interval is non negative.

(2) The total probability that the random variable assumes any value within the specified interval is one.

Thus, probability density function is used to determine probability that the value of random variable *X* lies within the specified interval *a* to *b* and it is denoted as $P(a < x < b)$. It is necessary to note here that probability for the definite value of continuous random variable *X* obtained by the probability density function is always zero (0) . i.e. $P(x = a) = 0$. Thus, the probabilities $P(a < x < b)$ and $P(a \le x \le b)$ obtained by using probability density function are always equal i.e. $P(a < x < b) = P(a \le x \le b)$.

Normal distribution is very important probability distribution among probability distributions for continuous random variable and is very useful distribution for higher statistical study. It can be defined as under:

If *X* is a random variable with mean μ and standard deviation σ and if its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; -\infty < x < \infty$$

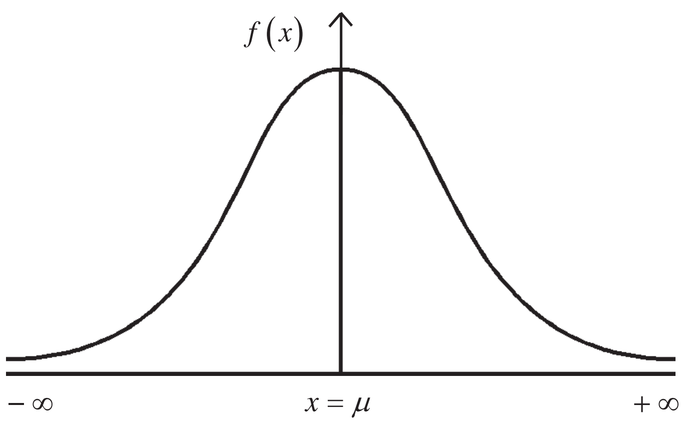
$$-\infty < \mu < \infty$$

$$0 < \sigma < \infty$$

where $\pi = 3.1416$ and $e = 2.7183$ are the constants

then *X* is called normal random variable and $f(x)$ is called probability density function of normal random variable. The distribution of this normal random variable *X* is called normal distribution and is denoted by $N(\mu, \sigma^2)$

A curve drawn by considering different values of normal random variable *X* and its respective values of probability density function $f(x)$ is called normal curve and is shown as under:



As shown in the above diagram, normal curve is completely bell shaped which shows that it is symmetric distribution.

3.2 Standard Normal Variable and Standard Normal Distribution

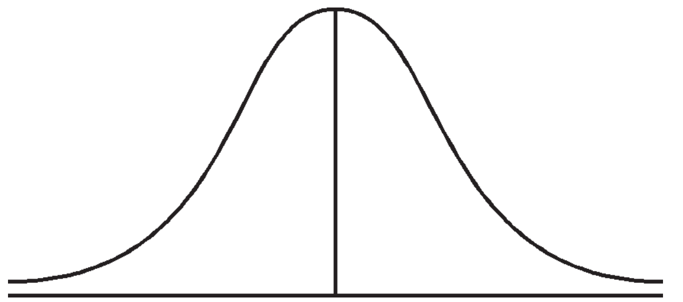
If X is a random normal variable with mean μ and standard deviation σ then random variable $Z = \frac{X-\mu}{\sigma}$ is called standard normal random variable and its probability density function is given below.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad ; -\infty < z < \infty$$

It can be seen here that probability density function of standard normal variable is a normal density function with mean zero (0) and standard deviation 1.

Note : During the further study of this chapter, we shall call normal variable X instead of normal random variable X and standard normal variable Z instead of standard normal random variable Z.

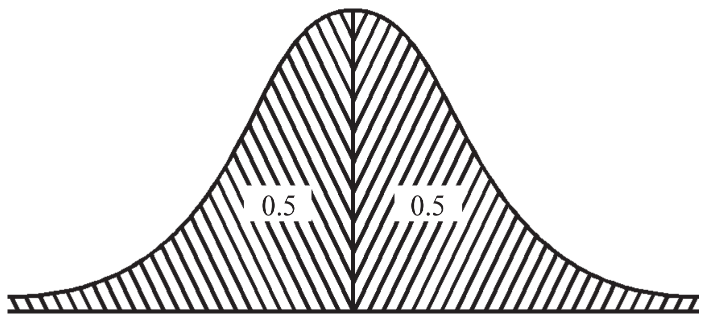
By plotting different values of standard normal variable Z and its respective values of $f(z)$ on graph paper, a completely bell shaped curve is obtained as under :



This curve is called standard normal curve and it is symmetrical to both the sides of $Z = 0$.

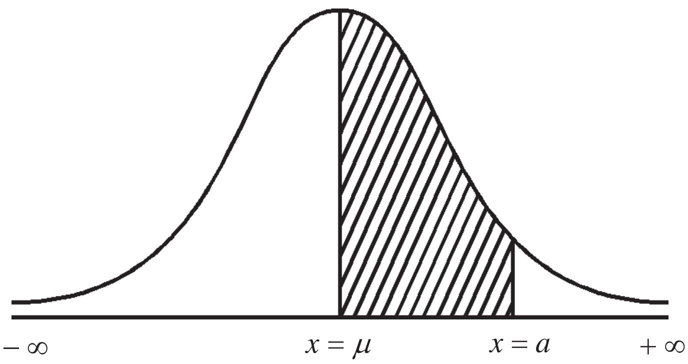
3.3 Method for Finding the Probability (area) from the Tables of Standard Normal Curve

We know that a normal curve is a curve of normal density function and it can be seen as under :



The area (probability) of the shaded region between the curve and X- axis is equal to 1. The normal curve of normal variable X is symmetrical about mean μ on both the sides and hence the perpendicular line at the point $X = \mu$ on X-axis divides the area (probability) of normal curve in two equal parts. The area (probability) to the right side of $X=\mu$ is 0.5 and is denoted by $P(X \geq \mu)=0.5$, whereas the area (probability) to the left side of $X=\mu$ is 0.5 and is denoted by $P(X \leq \mu)=0.5$.

In normal curve, probability that value of normal variable X lies between mean μ and its any specific value a $(a>\mu)$ can be shown by the area of shaded region between the x-axis and the perpendicular lines at $X=\mu$ and $X=a$. It can be shown as under :



In notation, this can be shown as $P(\mu \leq X \leq a)$.

For obtaining area under normal curve, first of all the normal variable X is changed into standard normal variable Z. By considering different positive values of standard normal variable Z, a table is prepared for obtaining area under normal curve for 0 to Z and by using this table the area can be obtained.

Note : A table for different values of standard normal variable is given on the last page of the book.

Suppose probability that a normal variable X assumes the value between mean μ and constant a $(a>\mu)$ is to be obtained then it is denoted as $P(\mu \leq X \leq a)$. Now, if standard deviation of normal variable X is σ,

When $X=\mu$ then $Z=\frac{X-\mu}{\sigma}=\frac{\mu-\mu}{\sigma}=\frac{0}{\sigma}=0$ and

When $X=a$ then $Z=\frac{X-\mu}{\sigma}=\frac{a-\mu}{\sigma}=Z_1$

Thus, $P(\mu \leq X \leq a)=P(0 \leq Z \leq Z_1)$

= area between $Z=0$ to $Z=Z_1$ obtained from tables of standard normal variable.

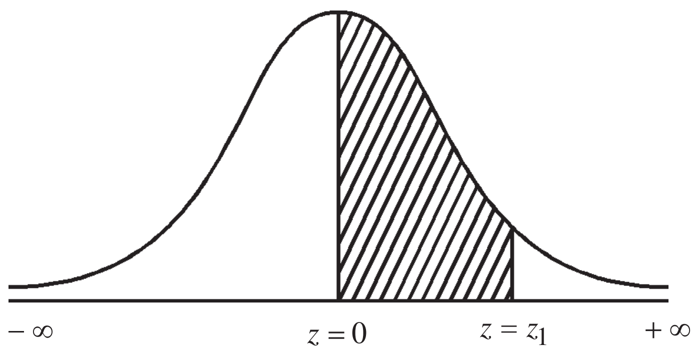


Illustration 1 : A normal distribution has mean 10 and standard deviation 2. Find the probabilities of (1) Normal variable *X* will take value between 10 and 12. (2) Normal variable *X* has the value between 8 and 10.

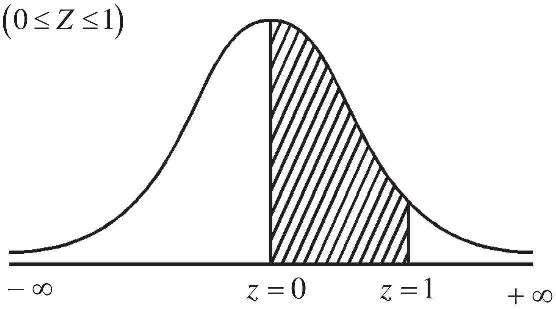
Here, mean $\mu = 10$ and standard deviation $\sigma = 2$.

(1) The probability that a normal variable X will take value between 10 and 12 is to be determined, i.e. to determine $P(10 \le X \le 12)$

$$\therefore P(10 \le X \le 12) = P\left(\frac{10-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{12-\mu}{\sigma}\right)$$

$$= P\left(\frac{10-10}{2} \le Z \le \frac{12-10}{2}\right)$$

$$= P(0 \le Z \le 1)$$



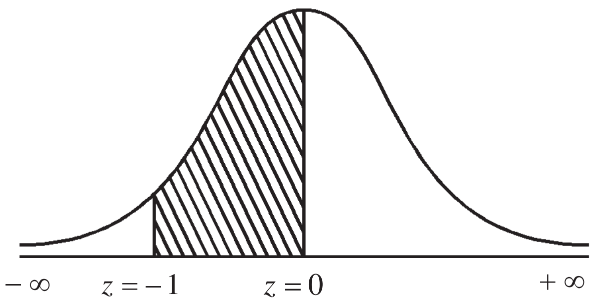
$= 0.3413$ (from the tables of standard normal variable)

(2) The probability that a normal variable X will take value between 8 and 10 is to be determined, i.e. to determine $P(8 \le X \le 10)$

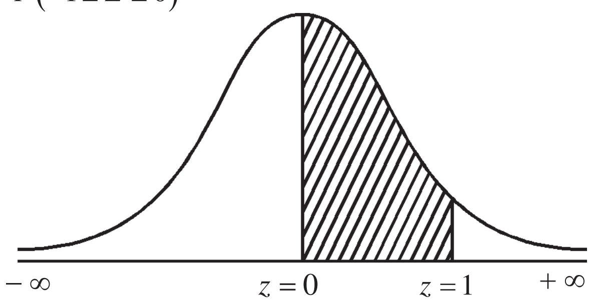
$$\therefore P(8 \le X \le 10) = P\left(\frac{8-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{10-\mu}{\sigma}\right)$$

$$= P\left(\frac{8-10}{2} \le Z \le \frac{10-10}{2}\right)$$

$$= P(-1 \le Z \le 0)$$



=



$= P(0 \le Z \le 1)$ ($\because$ Symmetry)

$= 0.3413$ (from the tables of standard normal variable)

Illustration 2 : A normal distribution has mean 20 and variance 16. Find the probabilities of (1) Normal variable *X* will take value less than 26 (2) Normal variable *X* has the value more than 14.

Here, mean $\mu = 20$ and variance $\sigma^2 = 16$.

$\therefore$ standard deviation $\sigma = 4$.

(1) The probability that a normal variable X will take value less than 26

$$= P(X \le 26)$$

$$= P\left(\frac{X-\mu}{\sigma} \le \frac{26-\mu}{4}\right)$$

$$= P\left(Z \le \frac{26-20}{4}\right)$$

$$= P(Z \le 1.5)$$