7. PROBABILITY DISTRIBUTIONS

Let us Learn

- Random variables
- Types of random variables
- Probability distribution of random variable.
  - Discrete random variable
  - Probability mass function
  - Expected values and variance
  - Continuous random variable
  - Probability density function
  - Cumulative distribution function

Let us Recall

- A random experiment and all possible outcomes of an experiment
- The sample space of a random experiment

Let us Study

7.1 Random variables :

We have already studied random experiments and sample spaces corresponding to random experiments. As an example, consider the experiment of tossing two fair coins. The sample space corresponding to this experiment contains four elements, namely {HH, HT,TH,TT}. We have already learnt to construct the sample space of any random experiment. However, the interest is not always in a random experiment and its sample space. We are often not interested in the outcomes of a random experiment, but only in some number obtained from the outcome. For example, in case of the experiment of tossing two fair coins, our interest may be only in the number of heads when two coins are tossed. In general, it is possible to associate a unique real number to every possible outcome of a random experiment. The number obtained from an outcome of a random experiment can take different values for different outcomes. This is why such a number is a variable. The value of this variable depends on the outcome of the random experiment, therefore it is called a random variable.

A random variable is usually denoted by capital letters like X, Y, Z, . . .

Consider the following examples to understand the concept of random variables.

(i) When we throw two dice, there are 36 possible outcomes, but if we are interested in the sum of the numbers on the two dice, then there are only 11 different possible values, from 2 to12.

(ii) If we toss a coin 10 times, then there are $2^{10} = 1024$ possible outcomes, but if we are interested in the number of heads among the 10 tosses of the coin, then there are only 11 different possible values, from 0 to 10.

(iii) In the experiment of randomly selecting four items from a lot of 20 items that contains 6 defective items, the interest is in the number of defective items among the selected four items. In this case, there are only 5 different possible outcomes, from 0 to 4.

In all the above examples, there is a rule to assign a unique value to every possible outcome of the random experiment. Since this number can change from one outcome to another, it is a variable. Also, since this number is obtained from outcomes of a random experiment, it is called a random variable.

A random variable is formally defined as follows.

Definition :

A random variable is a real-valued function defined on the sample space of a random experiment. In other words, the domain of a random variable is the sample space of a random experiment, while its co-domain is the set of real numbers.

Thus $X : S \rightarrow R$ is a random variable.

We often use the abbreviation r.v. to denote a random variable. Consider an experiment where three seeds are sown in order to find how many of them germinate. Every seed will either germinate or will not germinate. Let us use the letter Y when a seed germinates and the letter N when a seed does not germinate. The sample space of this experiment can then be written as

$S = \{YYY, YYN, YNY, NYY, YNN, NYN, NNY, NNN\}$, and n (S) = $2^3 = 8$.

None of these outcomes is a number. We shall try to represent every outcome by a number. Consider the number of times the letter Y appears in a possible outcome and denote it by X. Then, we have

$X(YYY) = 3$, $X(YYN) = X(YNY) = X(NYY) = 2$, $X(YNN) = X(NYN) = X(NNY) = 1$, $X(NNN) = 0$.

The variable X has four possible values, namely 0, 1, 2, and 3. The set of possible values of X is called the range of X. Thus, in this example, the range of X is the set $\{0, 1, 2, 3\}$.

A random variable is usually denoted by a capital letter, like X or Y. A particular value taken by the random variable is usually denoted by the small letter x. Note that x is always a real number and the set of all possible outcomes corresponding to a particular value x of X is denoted by the event $[X = x]$.

For example, in the experiment of three seeds, the random variable X takes four possible values, namely 0, 1, 2, 3. The four events are then defined as follows.

$$[X=0]=\{NNN\},$$

$$[X=1]=\{YNN, NYN, NNY\},$$

$$[X=2]=\{YYN, YNY, NYY\},$$

$$[X=3]=\{YYY\}.$$

Note that the sample space in this experiment is finite and so is the random variable defined on it.

A sample space need not always be finite. For example, the experiment of tossing a coin until a head is obtained. The sample space for this experiment is $S=\{H, TH, TTH, TTTH, \dots\}$.

Note that S contains an unending sequence of tosses required to get a head. Here, S is countably infinite.The random variable $X : S \to R$, denoting the number of tosses required to get a head, has the range $\{1, 2, 3, \dots\}$ which is also countably infinite.

7.2 Types of Random Variables :

There are two types of random variables, namely discrete and continuous.

7.2.1 Discrete Random Variables :

Definition : A random variable is said to be a discrete random variable if the number of its possible values is finite or countably infinite.

The values of a discrete random variable are usually denoted by non-negative integers, that is, $\{0, 1, 2, \dots\}$.

Examples of discrete random variables include the number of children in a family, the number of patients in a hospital ward, the number of cars sold by a dealer, number of stars in the sky and so on.

Note : The values of a discrete random variable are obtained by counting.

7.2.2 Continuous Random Variable :

Definition : A random variable is said to be a continuous random variable if the possible values of this random variable form an interval of real numbers.

A continuous random variable has uncountably infinite possible values and these values form an interval of real numbers.

Examples of continuous random variables include heights of trees in a forest, weights of students in a class, daily temperature of a city, speed of a vehicle, and so on.

The value of a continuous random variable is obtained by measurement. This value can be measured to any degree of accuracy, depending on the unit of measurement. This measurement can be represented by a point in an interval of real numbers.

The purpose of defining a random variable is to study its properties. The most important property of a random variable is its probability distribution. Many other properties of a random variable are obtained with help of its probability distribution. We shall now learn about the probability distribution of a random variable. We shall first learn the probability of a discrete random variable, and then learn the probability distribution of a continuous random variable.

7.3 Probability Distribution of Discrete Random Variables :

Let us consider the experiment of throwing two dice and noting the numbers on the upper-most faces of the two dice. The sample space of this experiment is

$S = \{(1, 1), (1, 2), ...(6, 6)\}$ and $n(S) = 36$.

Let X denote the sum of the two numbers in any single throw.

Then $\{2, 3, \cdots, 12\}$ is the set of possible values of X. Further,

$$[X = 2] = \{(1, 1)\},$$

$$[X = 3] = \{(1, 2), (2, 1)\},$$

$$\ldots$$

$$\ldots$$

$$[X = 12] = \{(6, 6)\}.$$

Next, all of these 36 possible outcomes are equally likely if the two dice are fair, that is, if each of the six faces have the same probability of being uppermost when the die is thrown.

As the result, each of these 36 possible outcomes has the same probability $= \frac{1}{36}$

This leads to the following results.

$$P[X = 2] = P\{(1, 1)\} = \frac{1}{36},$$

$$P[X = 3] = P\{(1, 2), (2, 1)\} = \frac{2}{36},$$

$$P[X = 4] = P\{(1, 3), (2, 2), (3, 1)\} = \frac{3}{36},$$

and so on.

The following table shows the probabilities of all possible values of X.

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Table 7.1

Such a description giving the values of the random variable X along with the corresponding probabilities is called the probability distribution of the random variable X.

In general, the probability distribution of a discrete random variable X is defined as follows.

Definition : The probability distribution of a discrete random variable X is defined by the following system of numbers. Let the possible values of X be denoted by $x_1, x_2, x_3, \ldots$, and the corresponding probabilities be denoted by $p_1, p_2, p_3, \ldots$, where $p_i = P[X = x_i]$ for $i = 1, 2, 3, \ldots$.

Note : A discrete random variable can have finite or infinite number of possible values, but they are countable.

Sometimes, the probability distribution of a discrete random variable is presented in the form of ordered pairs of the form $(x_1, p_1), (x_2, p_2), (x_3, p_3), \ldots$ A common practice is to present the probability distribution of a discrete random variable in a tabular form as shown below.

x_i	x_1	x_2	x_3	. . .
$P[X = x_i]$	p_1	p_2	p_3	. . .

Table 7.2

Note : If x_i is a possible value of X and $p_i = P[X = x_i]$, then there is an event $[E_i]$ in the sample space S such that $p_i = P[E_i]$. Since x_i is a possible value of X, $p_i = P[X = x_i] > 0$. Also, all possible values of X cover all sample points in the sample space S, and hence the sum of their probabilities is 1. That is, $p_i \geq 0$, for all i and $\sum_i p_i = 1$.

7.3.1 Probability Mass Function (p. m. f.) :

Sometimes the probability p_i of X taking the value x_i is a function of x_i for every possible value of X. Such a function is called the **probability mass function (p. m. f.)** of the discrete random variable X.

For example, consider the coin-tossing experiment where the random variable X is defined as the number of tosses required to get a head. Let probability of getting head be 't' and that of not getting head be $1 - t$. The possible values of X are given by the set of natural numbers, 1, 2, 3, . . . and

$$P[X = i] = (1 - t)^{i-1}\, t, \text{ for } i = 1, 2, 3, \ldots$$

This result can be verified by noting that if head is obtained for the first time on the i^{th} toss, then the first $i - 1$ tosses have resulted in tail. In other words, $X = i$ represents the event of having $i - 1$ tails followed by the first head on the toss.

We now define the probability mass function (p. m. f.) of a discrete random variable.

Definition : Let the possible values of a discrete random variable X be denoted by $x_1 , x_2 , x_3 , \ldots ,$ with the corresponding probabilities $p_i = P [X = x_i], i = 1, 2, \ldots$ the function p is called the probability mass function

(p. m. f.) of X, if (i) $p_i \geq 0$, i = 1,2,... n and (ii) $\sum_{i=1}^{n} p_i = 1$

For example, consider the experiment of tossing a coin 4 times and defining the random variable X as the number of heads in 4 tosses. The possible values of X are 0, 1, 2, 3, 4, and the probability distribution of X is given by the following table.

x	0	1	2	3	4
$P [X = x]$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Table 7.3

Note that : $$P [X = x] = {}^4C_x \left(\frac{1}{2}\right)^4, x = 0, 1, 2, 3, 4, \ldots$$

where ${}^4C_x$ is the number of ways of getting x heads in 4 tosses.

7.3.2 Cumulative Distribution Function (c. d. f.) :

The probability distribution of a discrete random variable can be specified with help of the p. m. f. It is sometimes more convenient to use the cumulative distribution function (c.d.f.) of the random variable. The cumulative distribution function (c. d. f.) of a discrete random variable is defined as follows.

Definition : The cumulative distribution function (c. d. f.) of a discrete random variable X is denoted by F and is defined as follows.

$$F (x) = P [X \leq x] = \sum_{x_i < x} P [X = x_i]$$

$$= \sum_{x_i < x} p_i$$

$$= \sum_{x_i < x} f (x_i)$$

where f is the probability mass function (p. m. f.) of the discrete random variable X.

For example, consider the experiment of tossing 4 coins and counting the number of heads.

We can form the next table for the probability distribution of X.

x	0	1	2	3	4
$f(x) = P[X = x]$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$
$F(x) = P[X \le x]$	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{11}{16}$	$\frac{15}{16}$	1

Table 7.4

For example, consider the experiment of tossing a coin till a head is obtained. The following table shows the p. m. f. and the c. d. f. of the random variable X, defined as the number of tosses required for the first head.

x	1	2	3	4	5	. . .
$f(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	. . .
$F(x)$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$	$\frac{31}{32}$	. . .

Table 7.5

It is possible to define several random variables on the same sample space. If two or more random variables are defined on the same sample space, their probability distributions need not be the same.

For example, consider the simple experiment of tossing a coin twice. The sample space of this experiment is $S = \{HH, HT, TH, TT\}$.

Let X denote the number of heads obtained in two tosses. Then X is a discrete random variable and its value for every outcome of the experiment is obtained as follows.

$$X(HH) = 2,\ X(HT) = X(TH) = 1,\ X(TT) = 0.$$

Let Y denote the number of heads minus the number of tails in two tosses. Then Y is also a discrete random variable and its value for every outcome of the experiment is obtained as follows.

$$Y(HH) = 2,\ Y(HT) = Y(TH) = 0,\ Y(TT) = -2.$$

Let $$Z = \frac{\text{Number of heads}}{\text{Number of tails} + 1}$$

Then Z is also a discrete random variable and its values for every outcome of the experiment is obtained as follows. $Z(HH) = 2,\ Z(HT) = Z(TH) = \frac{1}{2},\ Z(TT) = 0.$

These example show that it is possible to define many distinct random variables on the same sample space. Possible values of a discrete random variables can be positive or negative, integer or fraction.

SOLVED EXAMPLES

Ex. 1 : Two persons A and B play a game of tossing a coin thrice. If the result of a toss is head, A gets ₹ 2 from B. If the result of a toss is tail, B gets ₹ 1.5 from A. Let X denote the amount gained or lost by A. Show that X is a discrete random variable and show how it can be defined as a function on the sample space of the experiment.

Solution : X is a number whose value depends on the outcome of a random experiment.

Therefore, X is a random variable. Since the sample space of the experiment has only 8 possible outcomes, X is a discrete random variable. Now, the sample space of the experiment is

$$S = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}.$$

The values of X in rupees corresponding to these outcomes of the experiment are as follows.

$$X\,(\text{HHH}) = 2 \times 3 = ₹\ 6$$

$$X\,(\text{HHT}) = X\,(\text{HTH}) = X\,(\text{THH}) = 2 \times 2 - 1.50 \times 1 = ₹\ 2.50$$

$$X\,(\text{HTT}) = X\,(\text{THT}) = X\,(\text{TTH}) = 2 \times 1 - 1.50 \times 2 = ₹ - 1.00$$

$$X\,(\text{TTT}) = - 1.50 \times 3 = ₹ - 4.50$$

Here, a negative amount shows a loss to player A. This example shows that X takes a unique value for every element of the sample space and therefore X is a function on the sample space. Further, possible values of X are 4.50, 1, 2.50, 6.

Ex. 2 : A bag contains 1 red and 2 green balls. One ball is drawn from the bag at random, its colour is noted, and then ball is put back in the bag. One more ball is drawn from the bag at random and its colour is also noted. Let X denote the number of red balls drawn from the bag as described above. Derive the probability distribution of X.

Solution : Let the balls in the bag be denoted by r, g_1 , g_2 . The sample space of the experiment is then given by $$S = \{r\,r,\ r\,g_1,\ r\,g_2,\ g_1\,r,\ g_2\,r,\ g_1 g_1,\ g_1 g_2,\ g_2 g_1,\ g_2 g_2\}.$$

Since X is defined as the number of red balls, we have

$$X\,(\{r\,r\}) = 2,$$

$$X\,(\{r\,g_1\}) = X\,(r\,g_2) = X\,(g_1 r) = X\,(g_2 r) = 1,$$

$$X\,(\{g_1 g_1\}) = X\,(g_1 g_2) = X\,(g_2 g_1) = X\,(g_2 g_2) = 0.$$

Thus, X is a discrete random variable that can take values 0, 1, and 2.

The probability distribution of X is then obtained as follows :

x	0	1	2
$P\,[X = x\,]$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

Ex. 3 : Two cards are randomly drawn, with replacement, from a well shuffled deck of 52 playing cards. Find the probability distribution of the number of aces drawn.

Solution : Let X denote the number of aces among the two cards drawn with replacement.

Clearly, 0, 1, and 2 are the possible values of X. Since the draws are with replacement, the outcomes of the two draws are independent of each other. Also, since there are 4 aces in the deck of 52 cards,

$$\text{P [ace]} = \frac{4}{52} = \frac{1}{13} \text{ and P [non-ace]} = \frac{12}{13}\text{. Then}$$

$$\text{P } [X = 0] = \text{P [non-ace and non-ace]} = \frac{12}{13} \times \frac{12}{13} = \frac{144}{169},$$

$$\text{P } [X = 1] = \text{P [ace and non-ace]} + \text{P [non-ace and ace]}$$

$$= \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} = \frac{24}{169},$$

$$\text{and P } [X = 2] = \text{P [ace and ace]} = \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}.$$

The required probability distribution is then as follows.

x	0	1	2
$P[X = x]$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

Ex. 4 : A fair die is thrown. Let X denote the number of factors of the number on the upper face. Find the probability distribution of X.

Solution : The sample space of the experiment is $S = \{1, 2, 3, 4, 5, 6\}$. The values of X for the possible outcomes of the experiment are as follows.

$X(1) = 1$, $X(2) = 2$, $X(3) = 2$, $X(4) = 3$, $X(5) = 2$, $X(6) = 4$. Therefore,

$$p_1 = \text{P } [X = 1] = \text{P } [\{1\}] = \frac{1}{6}$$

$$p_2 = \text{P } [X = 2] = \text{P } [\{2, 3, 5\}] = \frac{3}{6}$$

$$p_3 = \text{P } [X = 3] = \text{P } [\{4\}] = \frac{1}{6}$$

$$p_4 = \text{P } [X = 4] = \text{P } [\{6\}] = \frac{1}{6}$$

The probability distribution of X is then as follows.

x	1	2	3	4
$P[X = x]$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Ex. 5 : Find the probability distribution of the number of doublets in three throws of a pair of dice.

Solution : Let X denote the number of doublets. Possible doublets in a pair of dice are (1, 1), (2, 2), (3, 3), (4, 4), (5, 5,), (6, 6).

Since the dice are thrown thrice, 0, 1, 2, and 3 are possible values of X. Probability of getting a doublet in a single throw of a pair of dice is $p = \frac{1}{6}$ and $q = 1 - \frac{1}{6} = \frac{5}{6}$.

$P[X = 0] = P[\text{ no doublet}] = qqq = \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \frac{125}{216}$.

$P[X = 1] = P[\text{one doublet}] = pqq + qpq + qqp = 3pq^2 = \frac{75}{216}$.

$P[X = 2] = P[\text{two doublets}] = ppq + pqp + qpp = 3p^2q = \frac{15}{216}$.

$P[X = 3] = \text{P}[\text{three doublets}] = ppp = \frac{1}{216}$.

Ex. 6 : The probability distribution of X is as follows :

x	0	1	2	3	4
$P[X = x]$	0·1	k	$2k$	$2k$	k

Find (i) k, (ii) P $[X < 2]$, (iii) P $[X \geq 3]$, (iv) P $[1 \leq X < 4]$, (v) F(2).

Solution : The table gives a probability distribution and therefore

P $[X = 0]$ + P $[X = 1]$ + P $[X = 2]$ + P $[X = 3]$ + P $[X = 4] = 1$.

That is, $0{\cdot}1 + k + 2k + 2k + k = 1$.

That is, $6k = 0{\cdot}9$. Therefore $k = 0{\cdot}15$.

(i) $k = 0{\cdot}15$.

(ii) P $[X < 2]$ = P $[X = 0]$ + P $[X = 1] = 0{\cdot}1 + k = 0{\cdot}1 + 0.15 = 0{\cdot}25$

(iii) P $[X \geq 3]$ = P $[X = 3]$ + P $[X = 4] = 2k + k = 3(0{\cdot}15) = 0{\cdot}45$

(iv) P $[1 \leq X < 4]$ = P $[X = 1]$ + P $[X = 2]$ + P $[X = 3]$ $= k + 2k + 2k = 5k$

$= 5(0{\cdot}15) = 0{\cdot}75$.

(v) F (2) = P $[X \leq 2]$ = P $[X = 0]$ + P $[X = 1]$ + P $[X = 2]$ $= 0{\cdot}1 + k + 2k = 0{\cdot}1 + 3k$

$= 0{\cdot}1 + 3(0{\cdot}15) = 0{\cdot}1 + 0{\cdot}45 = 0{\cdot}55$.

7.3.3 Expected value and Variance of a random variable :

In many problems, it is desirable to describe some feature of the random variable by means of a single number that can be computed from its probability distribution. Few such numbers are mean, median, mode and variance and standard deviation. In this section, we shall discuss mean and variance only. Mean is a measure of location or central tendency in the sense that it roughly locates a **middle** or **average value** of the random variable.

Definition : Let X be a random variable whose possible values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively. The expected value or arithmetic mean of X, denoted by E (X) or μ is defined by

$$E(X) = \mu = \sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + x_3 p_3 + \ldots + x_n p_n$$

In other words, the mean or expectation of a random variable X is the sum of the products of all possible values of X by their respective probabilities.

Definition : Let X be a random variable whose possible values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively. The variance of X, denoted by $Var\ (X)$ or σ_x^2 is defined as

$$\sigma_x^2 = Var\ (X) = \sum_{i=1}^{n} (x_i - \mu)^2 p_i$$

The non-negative number $\sigma_x = \sqrt{Var\ (X)}$ is called the **standard deviation** of the random variable X.

We can also use the simplified form of

$$Var\ (X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2$$

$$Var\ (X) = E(X^2) - [E(X)]^2 \quad \text{where} \sum_{i=1}^{n} x_i^2 p_i = E(X^2)$$

SOLVED EXAMPLES

Ex. 1 : Three coins are tossed simultaneously, X is the number of heads. Find expected value and variance of X.

Solution : $S = \{$HHH, HHT, HTH, THH, HTT, THT, TTH, TTT$\}$ and $X = \{0, 1, 2, 3\}$

$X = x_i$	$P = p_i$	$x_i p_i$	$x_i^2 p_i$
0	$\frac{1}{8}$	0	0
1	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{8}$
2	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{12}{8}$
3	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{9}{8}$
		$\sum_{i=1}^{n} x_i p_i = \frac{12}{8}$	$\sum_{i=1}^{n} x_i^2 p_i = \frac{24}{8}$

$$\text{Then E}(X) = \sum_{i=1}^{n} x_i p_i = \frac{12}{8} = 1{\cdot}5$$

$$Var\ (X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \frac{24}{8} - (1{\cdot}5)^2 = 3 - 2{\cdot}25 = 0{\cdot}75$$

Ex. 2 : Let a pair of dice be thrown and the random variable X be the sum of the numbers that appear on the two dice. Find the mean or expectation of X and variance of X.

Solution : The sample space of the experiment consists of 36 elementary events in the form of ordered pairs (x_i, y_i), where $x_i = 1, 2, 3, 4, 5, 6$ and $y_i = 1, 2, 3, 4, 5, 6$.

The random variable X i.e. the sum of the numbers on the two dice takes the values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12.

$X = x_i$	$P = p_i$	$x_i p_i$	$x_i^2 p_i$
2	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$
3	$\frac{2}{36}$	$\frac{6}{36}$	$\frac{18}{36}$
4	$\frac{3}{36}$	$\frac{12}{36}$	$\frac{48}{36}$
5	$\frac{4}{36}$	$\frac{20}{36}$	$\frac{100}{36}$
6	$\frac{5}{36}$	$\frac{30}{36}$	$\frac{180}{36}$
7	$\frac{6}{36}$	$\frac{42}{36}$	$\frac{294}{36}$
8	$\frac{5}{36}$	$\frac{40}{36}$	$\frac{320}{36}$
9	$\frac{4}{36}$	$\frac{36}{36}$	$\frac{324}{36}$
10	$\frac{3}{36}$	$\frac{30}{36}$	$\frac{300}{36}$
11	$\frac{2}{36}$	$\frac{22}{36}$	$\frac{242}{36}$
12	$\frac{1}{36}$	$\frac{12}{36}$	$\frac{144}{36}$
		$\sum_{i=1}^{n} x_i p_i = \frac{252}{36} = 7$	$\sum_{i=1}^{n} x_i^2 p_i = \frac{1974}{36} = 54{\cdot}83$

$$\text{Then } E(X) = \sum_{i=1}^{n} x_i p_i = 7$$

$$Var(X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = 54{\cdot}83 - (7)^2$$

$$= 54{\cdot}83 - 49$$

$$= 5{\cdot}83$$

Ex. 3 : Find the mean and variance of the number randomly selected from 1 to 15.

Solution : The sample space of the experiment is $S = \{1, 2, 3, \ldots, 15\}$.

Let X denote the number selected.

Then X is a random variable which can take values 1, 2, 3, …, 15. Each number selected is equiprobable therefore

$$P(1) = P(2) = P(3) = \ldots = P(15) = \frac{1}{15}$$

$$\mu = E(X) = \sum_{i=1}^{n} x_i p_i = 1 \times \frac{1}{15} + 2 \times \frac{1}{15} + 3 \times \frac{1}{15} + \ldots + 15 \times \frac{1}{15}$$

$$= (1 + 2 + 3 + \ldots + 15) \times \frac{1}{15} = \left(\frac{15 \times 16}{2}\right) \times \frac{1}{15} = 8$$

$$Var(X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = 1^2 \times \frac{1}{15} + 2^2 \times \frac{1}{15} + 3^2 \times \frac{1}{15} + \ldots + 15^2 \times \frac{1}{15} - (8)^2$$

$$= (1^2 + 2^2 + 3^2 + \ldots + 15^2) \times \frac{1}{15} - (8)^2$$

$$= \left(\frac{15 \times 16 \times 31}{6}\right) \times \frac{1}{15} - (8)^2$$

$$= 82{\cdot}67 - 64 = 18{\cdot}67$$

Ex. 4 : Two cards are drawn simultaneously (or successively without replacement) from a well shuffled pack of 52 cards. Find the mean, variance and standard deviation of the number of kings drawn.

Solution : Let X denote the number of kings in a draw of two cards. X is a random variable which can assume the values 0, 1 or 2.

Then $P(X = 0) = P$ (no card is king) $= \frac{^{48}C_2}{^{52}C_2} = \frac{48 \times 47}{52 \times 51} = \frac{188}{221}$

Then $P(X = 1) = P$ (exactly one card is king) $= \frac{^{4}C_1 \times {}^{48}C_1}{^{52}C_2} = \frac{4 \times 48 \times 27}{52 \times 51} = \frac{32}{221}$

Then $P(X = 2) = P$ (both cards are king) $= \frac{^{4}C_2}{^{52}C_2} = \frac{4 \times 3}{52 \times 51} = \frac{1}{221}$

$$\mu = E(X) = \sum_{i=1}^{n} x_i p_i = 0 \times \frac{188}{221} + 1 \times \frac{32}{221} + 2 \times \frac{1}{221} = \frac{34}{221}$$

$$Var(X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \left(0^2 \times \frac{188}{221} + 1^2 \times \frac{32}{221} + 2^2 \times \frac{1}{221}\right) - \left(\frac{34}{221}\right)^2$$

$$= \frac{36}{221} - \frac{1156}{48841} = \frac{6800}{48841} = 0{\cdot}1392$$

$$\sigma = \sqrt{Var(X)} = \sqrt{0{\cdot}1392}$$

EXERCISE 7.1

1. Let X represent the difference between number of heads and number of tails obtained when a coin is tossed 6 times. What are possible values of X ?

2. An urn contains 5 red and 2 black balls. Two balls are drawn at random. X denotes number of black balls drawn. What are possible values of X ?

3. State which of the following are not the probability mass function of a random variable. Give reasons for your answer.

(i)

X	0	1	2
$P(X)$	0·4	0·4	0·2

(ii)

X	0	1	2	3	4
$P(X)$	0·1	0·5	0·2	–0·1	0·2

(iii)

X	0	1	2
$P(X)$	0·1	0·6	0·3

(iv)

Z	3	2	1	0	–1
$P(Z)$	0·3	0·2	0·4	0	0·05

(v)

Y	–1	0	1
$P(Y)$	0·6	0·1	0·2

(vi)

X	0	–1	–2
$P(X)$	0·3	0·4	0·3

4 Find the probability distribution of (i) number of heads in two tosses of a coin. (ii) Number of tails in the simultaneous tosses of three coins. (iii) Number of heads in four tosses of a coin.

5. Find the probability distribution of the number of successes in two tosses of a die, where a success is defined as number greater than 4 appears on at least one die.

6. From a lot of 30 bulbs which include 6 defectives, a sample of 4 bulbs is drawn at random with replacement. Find the probability distribution of the number of defective bulbs.

7. A coin is biased so that the head is 3 times as likely to occur as tail. If the coin is tossed twice, find the probability distribution of number of tails.

8. A random variable X has the following probability distribution :

X	0	1	2	3	4	5	6	7
$P(X)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Determine : (i) k (ii) $P(X < 3)$ (iii) $P(X > 4)$

9. Find expected value and variance of X for the following p.m.f.

X	–2	–1	0	1	2
$P(X)$	0·2	0·3	0·1	0·15	0·25

10. Find expected value and variance of X ,where X is number obtained on uppermost face when a fair die is thrown.

11. Find the mean number of heads in three tosses of a fair coin.

12. Two dice are thrown simultaneously. If X denotes the number of sixes, find the expectation of X.

13. Two numbers are selected at random (without replacement) from the first six positive integers. Let X denote the larger of the two numbers obtained. Find E (X).

14. Let X denote the sum of the numbers obtained when two fair dice are rolled. Find the standard deviation of X.

15. A class has 15 students whose ages are 14, 17, 15, 14, 21, 17, 19, 20, 16, 18, 20, 17, 16, 19 and 20 years. One student is selected in such a manner that each has the same chance of being chosen and the age X of the selected student is recorded. What is the probability distribution of the random variable X ? Find mean, variance and standard deviation of X.

16. In a meeting, 70% of the members favour and 30% oppose a certain proposal. A member is selected at random and we take $X = 0$ if he opposed, and $X = 1$ if he is in favour. Find E (X) and *Var* (X).

7.4 Probability Distribution of a Continuous Random Variable :

A continuous random variable differs from a discrete random variable in the sense that the possible values of a continuous random variable form an interval of real numbers. In other words, a continuous random variable has uncountably infinite possible values.

For example, the time an athlete takes to complete a thousand-meter race is a continuous random variable.

We shall extend what we learnt about a discrete random variable to a continuous random variable. More specifically, we shall study the probability distribution of a continuous random variable with help of its probability density function (p. d. f.) and its cumulative distribution function (c. d. f.). If the possible values of a continuous random variable X form the interval $[a, b]$, where a and b are real numbers and $a < b$, then the interval $[a, b]$ is called the support of the continuous random variable X. The support of a continuous random variable is often denoted by S.

In case of a discrete random variable X that takes finite or countably infinite distinct values, the probability $P\,[X = x]$ is determined for every possible value x of the random variable X. The probability distribution of a continuous random variable is not defined in terms of probabilities of possible values of the random variable since the number of possible values are unaccountably infinite. Instead, the probability distribution of a continuous random variable is characterized by probabilities of intervals of the form $[\,c, d\,]$, where $c < d$. That is, for a continuous random variable, the interest is in probabilities of the form $P\,[c < X < d]$, where $a \le c < d \le b$.

This probability is obtained by integrating a function of X over the interval $[c, d]$. Let us first define the probability density function (p.d.f.) of a continuous random variable.

7.4.1 Probability Density Function (p. d. f.) :

Let X be a continuous random variable defined on interval $S = (a, b)$. A non-negative integrable function $f(x)$ is called the probability density function (p. d. f.) of X if it satisfies the following conditions.

1. $f(x)$ is positive or zero every where in S, that is, $f(x) \geq 0$, for all $x \in S$.

2. The area under the curve $y = f(x)$ over S is 1. That is, $\int_S f(x)\, dx = 1$

3. The probability that X takes a value in A, where A is some interval, is given by the integral of $f(x)$ over that interval. That is

$$P[X \in A] = \int_A f(x)\, dx$$

7.4.2 Cumulative Distribution Functions (c. d. f.) :

The cumulative distribution function for continuous random variables is just a straightforward extension of that of the discrete case. All we need to do is replace the summation with an integral.

Definition : The **cumulative distribution function** (c. d. f.) of a continuous random variable X is defined as :

$$F(x) = \int_a^x f(t)\, dt \qquad \text{for } a < x < b.$$

You might recall, for discrete random variables, that $F(x)$ is, in general, a non-decreasing step function. For continuous random variables, $F(x)$ is a non-decreasing **continuous function**.

SOLVED EXAMPLES

Ex. 1 : Let X be a continuous random variable whose probability density function is $f(x) = 3x^2$, for $0 < x < 1$. note that $f(x)$ is not $P[X = x]$.

For example, $f(0.9) = 3(0.9)^2 = 2{\cdot}43 > 1$, which is clearly not a probability. In the continuous case, $f(x)$ is the height of the curve at $X = x$, so that the total area under the curve is 1. Here it is areas under the curve that define the probabilities.

Solution : Now, let's start by verifying that $f(x)$ is a valid probability density function.

For this, note the following results.

1. $f(x) = 3x^2 \geq 0$ for all $x \in [0, 1]$.

2. $\int_0^1 f(x) = \int_0^1 3x^2\, dx = 1$

Therefore, the function $f(x) = 3x^2$, for $0 < x < 1$ is a proper probability density function.

Also, for real numbers c and d such that $0 \le c < d \le 1$, note that

$$P[c < X < d] = \int_c^d f(x)\, dx = \int_c^d 3x^2\, dx = \left[x^3\right]_c^d = d^3 - c^3 > 0$$

- What is the probability that X falls between $\frac{1}{2}$ and 1? That is, what is $P\left[\frac{1}{2} < X < 1\right]$?

Substitute $c = \frac{1}{2}$ and $d = 1$ in the above integral to obtain

$$P\left[\frac{1}{2} < X < 1\right] = 1^3 - \left(\frac{1}{2}\right)^3 = 1 - \frac{1}{8} = \frac{7}{8}.$$

- What is $P\left(X = \frac{1}{2}\right)$?

See that the probability is 0 . This is so because

$$\int_c^d f(x)\, dx = \int_{1/2}^{1/2} x^3\, dx = 1 - 1 = 0.$$

$\left[\text{The ordinate AB, with A}\left(\frac{1}{2}, 0\right) \text{and B}\left(\frac{1}{2}, \frac{1}{8}\right) \text{is degenerate case of rectangle and has area 0}\right]$

As a matter of fact, in general, if X is a continuous random variable, then the probability that X takes any specific value x is 0. That is, when X is a continuous random variable, then

$P[X = x] = 0$ for every x in the support.

Ex. 2 : Let X be a continuous random variable whose probability density function is $f(x) = \frac{x^3}{4}$ for an interval $0 < x < c$. What is the value of the constant c that makes $f(x)$ a valid probability density function?

Solution : Note that the integral of the p. d. f. over the support of the random variable must be

That is, $\int_0^c f(x)\, dx = 1$.

That is, $\int_0^c \left(\frac{x^3}{4}\right) dx = 1$. But, $\int_0^c \left(\frac{x^3}{4}\right) dx = \left[\frac{x^4}{16}\right]_0^c = \frac{c^4}{16}$. Since this integral must be equal to 1,

we have $\frac{c^4}{16} = 1$, or equivalently $c^4 = 16$, so that $c = 2$ since c must be positive.

Ex. 3 : Let's return to the example in which X has the following probability density function :

$$f(x) = 3x^2$$

for $0 < x < 1$. What is the cumulative distribution function $F(x)$?

Solution : $F(x) = \int_{-0}^{x} f(x)\, dx = \int_0^x 3x^2\, dx = \left[x^3\right]_0^x = x^3$

Ex. 4 : Let's return to the example in which X has the following probability density function :

$f(x) = \dfrac{x^3}{4}$ for $0 < x < 4$. What is the cumulative distribution function X ?

Solution : $F(x) = \int_0^x f(x)\,dx = \int_0^x \dfrac{x^3}{4}\,dx = \dfrac{1}{4}\left[\dfrac{x^4}{4}\right]_0^x = \dfrac{1}{16} = [x^4 - 0] = \dfrac{x^4}{16}$

Ex. 5 : Suppose the p.d.f. of a continuous random variable X is defined as:

$f(x) = x + 1$, for $-1 < x < 0$, and $f(x) = 1 - x$, for $0 \leq x < 1$.

Find the c.d.f. $F(x)$.

Solution : If we look the p.d.f. it is defined in two steps

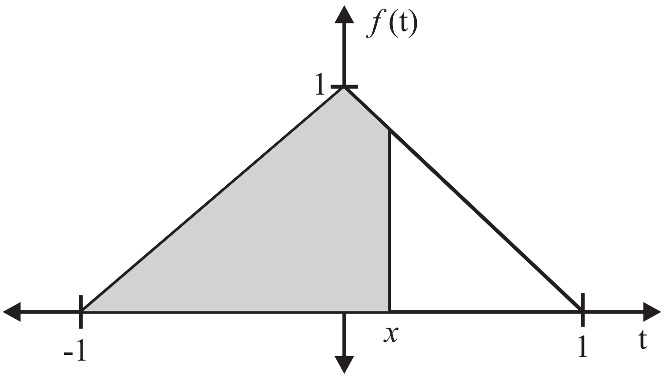

Fig. 7.1

Now for the other two intervals :

For $-1 < x < 0$ and $0 < x < 1$.

$$F(x) = \int_{-1}^{x} (x+1)\,dx$$

$$= \left[\frac{x^2}{2} + x\right]_{-1}^{x} = \left(\frac{x^2}{2} + x\right) - \left(\frac{1}{2} - 1\right)$$

$$= \frac{x^2}{2} + x + \frac{1}{2} = \frac{x^2 + 2x + 2}{2}$$

$$= \frac{(x+1)^2}{2}$$

$$F(0) = \frac{1}{2}$$

For $0 < x < 1$

$$F(x) = P(-1 < x < 1) = P(-1 < x < 0) + P(0 < x < 1) = \int_{-1}^{0} (x+1)\,dx + \int_0^x (1-x)\,dx$$

$$= \frac{1}{2} + \int_0^x (1-x)\,dx$$

$$= \frac{1}{2} + x - \frac{x^2}{2}$$

$F(x) = 0$, for $x \leq -1$

$F(x) = \dfrac{1}{2}(x+1)^2$, for $-1 < x \leq 0$

$= \dfrac{1}{2} + x - \dfrac{x^2}{2}$, for $0 < x < 1$

$=1$, for $x \geq 1$

If probability function $f(x)$ is defined on (a, b) with $f(x) \geq 0$ and $\int_a^b f(x)\,dx = 1$, then we can extend this function to the whole of R as follows.

For $x \leq a$ and $x \geq b$, define $f(x) = 0$.

Then note that $\int_{-\infty}^{t} f(x) = 0$, for $t \leq a$ and for $x \geq b$

$$\int_{-\infty}^{t} f(x)\,dx = \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{b} f(x)\,dx + \int_{b}^{t} f(x)\,dx = 0 + 1 + 0$$

Thus $F(t) = 0$, for $t \leq a$ and $F(t) = 1$, for $t \geq b$

Ex. 6 : Verify if the following functions are p.d.f. of a continuous r.v. X.

(i) $f(x) = e^{-x}$, for $0 < x < \infty$ and $= 0$, otherwise.

(ii) $f(x) = \dfrac{x}{2}$, for $-2 < x < 2$ and $= 0$, otherwise.

Solution : (i) e^{-x} is ≥ 0 for any value of x since $e > 0$,

$\therefore \quad e^{-x} > 0$, for $0 < x < \infty$

$$\int_0^\infty f(x)\,dx = \int_0^\infty e^{-x}\,dx = \left[-e^{-x}\right]_0^\infty = \left[\frac{1}{e^\infty} - e^0\right] = -(0-1) = 1$$

Both the conditions of p.d.f. are satisfied $f(x)$ is p.d.f. of r.v.

(ii) $f(x) < 0$ i.e. negative for $-2 < x < 0$ therefore f(x) is not p.d.f.

Ex. 7 : Find k if the following function is the p.d.f. of r.v. X.

$f(x) = kx^2(1-x)$, for $0 < x < 1$ and $= 0$, otherwise.

Solution : Since $f(x)$ is the p.d.f. of r.v. X

$$\int_0^1 kx^2(1-x)\,dx = 1$$

$$\therefore \quad \int_0^1 k(x^2 - x^3)\,dx = 1$$

$$\therefore \quad k\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 = 1$$

$$\therefore \quad k\left\{\left[\frac{1}{3} - \frac{1}{4}\right] - (0)\right\} = 1$$

$$\therefore \quad k \times \frac{1}{12} = 1 \qquad \therefore \quad k = 12$$

Ex. 8 : For each of the following p.d.f. of r.v. X, find (a) $P(X < 1)$ and (b) $P(|X| < 1)$

(i) $f(x) = \dfrac{x^2}{18}$, for $-3 < x < 3$ and $= 0$, otherwise.

(ii) $f(x) = \dfrac{x+2}{18}$, for $-2 < x < 4$ and $= 0$, otherwise.

Solution :

(i) (a) $P(X < 1) \quad = \displaystyle\int_{-3}^{1} \frac{x^2}{18}\,dx = \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-3}^{1} = \frac{1}{54}\left[1-(-3)^3\right] = \frac{1}{54}(1+27) = \frac{28}{54} = \frac{14}{27}$

(b) $P(|X| < 1) \quad = P(-1 < x < 1) = \displaystyle\int_{-1}^{1} \frac{x^2}{18}\,dx = \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-1}^{1}$

$$= \frac{1}{54}\left[1-(-1)^3\right] \quad = \frac{1}{54}(1+1) = \frac{2}{54} = \frac{1}{27}$$

(ii) (a) $P(X < 1) \quad = \displaystyle\int_{-2}^{1} \frac{x+2}{18}\,dx = \frac{1}{18}\left[\frac{x^2}{2} + 2x\right]_{-2}^{1}$

$$= \frac{1}{18}\left\{\left(\frac{1}{2}+2\right) - \left(\frac{(-2)^2}{2} + 2(-2)\right)\right\} = \frac{1}{18}\left\{\frac{5}{2}+2\right\} = \frac{1}{18} \times \frac{9}{2} = \frac{1}{4}$$

(b) $P(|X| < 1) \quad = P(-1 < x < 1) = \displaystyle\int_{-1}^{1} \frac{x+2}{18}\,dx = \frac{1}{18}\left[\frac{x^2}{2} + 2x\right]_{-1}^{1}$

$$= \frac{1}{18}\left\{\left(\frac{1}{2}+2\right) - \left(\frac{1}{2}-2\right)\right\} = \frac{1}{18}\left\{\frac{5}{2}+\frac{3}{2}\right\} = \frac{1}{18} \times 4 = \frac{2}{9}$$

Ex. 9 : Find the c.d.f. $F(x)$ associated with p.d.f. $f(x)$ of r.v. X where

$f(x) = 3(1 - 2x^2)$ for $0 < x < 1$ and $= 0$, otherwise.

Solution : Since $f(x)$ is p.d.f. of r.v. therefore c.d.f. is

$$F(x) = \int_0^x 3(1-2x^2)\,dx = \left[3\left(x - \frac{2x^3}{3}\right)\right]_0^x = \left[3x - 2x^3\right] = 3x - 2x^3$$

EXERCISE 7.2

1. Verify which of the following is p.d.f. of r.v. X :

(i) $f(x) = \sin x$, for $0 \le x \le \dfrac{\pi}{2}$ (ii) $f(x) = x$, for $0 \le x \le 1$ and $= 2 - x$ for $1 < x < 2$

(iii) $f(x) = 2$, for $0 \le x \le 1$

2. The following is the p.d.f. of r.v. $X : f(x) = \frac{x}{8}$, for $0 < x < 4$ and $= 0$ otherwise

Find (a) $P(x < 1{\cdot}5)$ (b) $P(1 < x < 2)$ (c) $P(x > 2)$

3. It is known that error in measurement of reaction temperature (in 0° c) in a certain experiment is continuous r.v. given by

$f(x) = \frac{x^2}{3}$, for $-1 < x < 2$ and $= 0$ otherwise

(i) Verify whether $f(x)$ is p.d.f. of r.v. X. (ii) Find $P(0 < x \leq 1)$

(iii) Find probability that X is negative.

4. Find k if the following function represent p.d.f. of r.v. X.

(i) $f(x) = kx$, for $0 < x < 2$ and $= 0$ otherwise, Also find $P\left(\frac{1}{4} < x < \frac{3}{2}\right)$.

(ii) $f(x) = kx(1 - x)$, for $0 < x < 1$ and $= 0$ otherwise, Also find $P\left(\frac{1}{4} < x < \frac{1}{2}\right)$, $P\left(x < \frac{1}{2}\right)$.

5. Let X be amount of time for which a book is taken out of library by randomly selected student and suppose X has p.d.f.

$f(x) = 0{\cdot}5x$, for $0 \leq x \leq 2$ and $= 0$ otherwise.

Calculate: (i) $P(X \leq 1)$ (ii) $P(0{\cdot}5 \leq x \leq 1{\cdot}5)$ (iii) $P(x \geq 1{\cdot}5)$

6. Suppose that X is waiting time in minutes for a bus and its p.d.f. is given by

$f(x) = \frac{1}{5}$, for $0 \leq x \leq 5$ and $= 0$ otherwise.

Find the probability that (i) waiting time is between 1 and 3

(ii) waiting time is more than 4 minutes.

7. Suppose error involved in making a certain measurement is continuous r.v. X with p.d.f.

$f(x) = k(4 - x^2)$, for $-2 \leq x \leq 2$ and $= 0$ otherwise.

Compute : (i) $P(x > 0)$ (ii) $P(-1 < x < 1)$ (iii) $P(-0{\cdot}5 < x$ or $x > 0{\cdot}5)$

8. The following is the p.d.f. of continuous r.v.

$f(x) = \frac{x}{8}$, for $0 < x < 4$ and $= 0$ otherwise.

(i) Find expression for c.d.f. of X (ii) Find $F(x)$ at $x = 0{\cdot}5$, 1.7 and 5.

9. Given the p.d.f. of a continuous r.v. X , $f(x) = \frac{x^2}{3}$, for $-1 < x < 2$ and $= 0$ otherwise

Determine c.d.f. of X hence find $P(x < 1)$, $P(x < -2)$, $P(X > 0)$, $P(1 < x < 2)$

10. If a r.v. X has p.d.f.,

$f(x) = \frac{c}{x}$, for $1 < x < 3$, $c > 0$, Find c, E(X) and Var (X).

Let us Remember

- A random variable (r.v.) is a real-valued function defined on the sample space of a random experiment.

  The domain of a random variable is the sample space of a random experiment, while its co-domain is the real line.

  Thus $X : S \to R$ is a random variable.

  There are two types of random variables, namely discrete and continuous.

- **Discrete random variable :** Let the possible values of discrete random variable X be denoted by $x_1, x_2, x_3, \ldots$, and the corresponding probabilities be denoted by $p_1, p_2, p_3, \ldots$, where $p_i = P[X = x_i]$ for $i = 1, 2, 3, \ldots$. If there is a function f such that $p_i = P[X = x_i] = f(x_i)$ for all possible values of X, then f is called the probability mass function (p. m. f.) of X.

  Note : If x_i is a possible value of X and $p_i = P[X = x_i]$, then there is an event E_i in the sample space S such that $p_i = P[E_i]$. Since x_i is a possible value of X, $p_i = P[X = x_i] > 0$. Also, all possible values of X cover all sample points in the sample space S, and hence the sum of their probabilities is 1. That is, $p_i > 0$ for all i and $\sum p_i = 1$.

- **c.d.f ($F(x)$) :** The cumulative distribution function (c. d. f.) of a discrete random variable X is denoted by F and is defined as follows.

$$F(x) = P[X \le x] \qquad = \sum_{x_i < x} P[X = x_i]$$

$$= \sum_{x_i < x} p_i$$

$$= \sum_{x_i < x} f(x_i)$$

- **Expected Value or Mean of Discrete r. v. :** Let X be a random variable whose possible values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively. The expected value or arithmetic mean of X, denoted by $E(X)$ or μ is defined by

$$E(X) = \mu = \sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + x_3 p_3 + \ldots + x_n p_n$$

  In other words, the mean or expectation of a random variable X is the sum of the products of all possible values of X by their respective probabilities.

Variance of Discrete r. v. : Let X be a random variable whose possible values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively. The variance of X, denoted by $Var(X)$ or σx^2 is defined as

$$\sigma_x^2 = Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p_i$$

The non-negative number $\sigma_x = \sqrt{Var(X)}$

is called the **standard deviation** of the random variable X.

Another formula to find the variance of a random variable. We can also use the simplified form of

$$Var(X) = \left(\sum_{i=1}^{n} x_i^2 p_i\right) - \left(\sum_{i=1}^{n} x_i p_i\right)^2$$

$$Var(X) = E(X^2) - [E(X)]^2 \qquad \text{where } \sum_{i=1}^{n} x_i^2 p_i = E(X^2)$$

Probability Density Function (p. d. f.) : Let X be a continuous random variable defined on interval $S = (a, b)$. A non-negative integrable function $f(x)$ is called the probability density function (p. d. f.) of X if it satisfies the following conditions.

1. $f(x)$ is positive every where in S, that is, $f(x) > 0$, for all $x \in S$.
2. The area under the curve $f(x)$ over S is 1. That is, $\int_S f(x)\, dx = 1$
3. The probability that X takes a value in A, where A is some interval, is given by the integral of $f(x)$ over that interval. That is

$$P[X \in A] = \int_A f(x)\, dx$$

The **cumulative distribution function** (c. d. f.) of a continuous random variable X is defined as :

$$F(x) = \int_a^x f(t)\, dt \qquad \text{for } a < x < b.$$

MISCELLANEOUS EXERCISE 7

(I) Choose the correct option from the given alternatives :

(1) P.d.f. of a.c.r.v X is $f(x) = 6x(1 - x)$, for $0 \le x \le 1$ and $= 0$, otherwise (elsewhere)

If $P(X < a) = P(X > a)$, then $a =$

(A) 1 (B) $\frac{1}{2}$ (C) $\frac{1}{3}$ (D) $\frac{1}{4}$

(2) If the p.d.f of a.c.r.v. X is $f(x) = 3(1 - 2x^2)$, for $0 < x < 1$ and $= 0$, otherwise (elsewhere) then the c.d.f of X is $F(x) =$

(A) $2x - 3x^2$ (B) $3x - 4x^3$ (C) $3x - 2x^3$ (D) $2x^3 - 3x$

(3) If the p.d.f of a.c.r.v. X is $f(x) = \dfrac{x^2}{18}$, for $-3 < x < 3$ and $= 0$, otherwise then $P(|X| < 1) =$

(A) $\dfrac{1}{27}$ (B) $\dfrac{1}{28}$ (C) $\dfrac{1}{29}$ (D) $\dfrac{1}{26}$

(4) If a d.r.v. X takes values 0, 1, 2, 3, . . . which probability $P(X = x) = k(x + 1)\cdot 5^{-x}$, where k is a constant, then $P(X = 0) =$

(A) $\dfrac{7}{25}$ (B) $\dfrac{16}{25}$ (C) $\dfrac{18}{25}$ (D) $\dfrac{19}{25}$

(5) If p.m.f. of a d.r.v. X is $P(X = x) = \dfrac{({}^5C_x)}{2^5}$, for $x = 0, 1, 2, 3, 4, 5$ and $= 0$, otherwise

If $a = P(X \leq 2)$ and $b = P(X \geq 3)$, then $E(X) =$

(A) $a < b$ (B) $a > b$ (C) $a = b$ (D) $a + b$

(6) If p.m.f. of a d.r.v. X is $P(X = x) = \dfrac{x}{n(n+1)}$, for $x = 1, 2, 3, . . ., n$ and $= 0$, otherwise then $E(X) =$

(A) $\dfrac{n}{1} + \dfrac{1}{2}$ (B) $\dfrac{n}{3} + \dfrac{1}{6}$ (C) $\dfrac{n}{2} + \dfrac{1}{5}$ (D) $\dfrac{n}{1} + \dfrac{1}{3}$

(7) If p.m.f. of a d.r.v. X is $P(x) = \dfrac{c}{x^3}$, for $x = 1, 2, 3$ and $= 0$, otherwise (elsewhere) then $E(X) =$

(A) $\dfrac{343}{297}$ (B) $\dfrac{294}{251}$ (C) $\dfrac{297}{294}$ (D) $\dfrac{294}{297}$

(8) If the a d.r.v. X has the following probability distribution :

X	-2	-1	0	1	2	3
$P(X = x)$	0.1	k	0.2	$2k$	0.3	k

then $P(X = -1) =$

(A) $\dfrac{1}{10}$ (B) $\dfrac{2}{10}$ (C) $\dfrac{3}{10}$ (D) $\dfrac{4}{10}$

(9) If the a d.r.v. X has the following probability distribution :

X	1	2	3	4	5	6	7
$P(X = x)$	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

then $k =$

(A) $\dfrac{1}{7}$ (B) $\dfrac{1}{8}$ (C) $\dfrac{1}{9}$ (D) $\dfrac{1}{10}$

(10) The expected value of X for the following p.m.f.

X	-2	-1	0	1	2
$P(X)$	0·3	0·3	0·1	0·05	0·25

(A) 0·85 (B) – 0·35 (C) 0·15 (D) – 0·15

(II) Solve the following :

(1) Identify the random variable as either discrete or continuous in each of the following. Write down the range of it.

(i) An economist is interested the number of unemployed graduate in the town of population 1 lakh.

(ii) Amount of syrup prescribed by physician.

(iii) The person on the high protein diet is interested gain of weight in a week.

(iv) 20 white rats are available for an experiment. Twelve rats are male. Scientist randomly selects 5 rats number of female rats selected on a specific day.

(v) A highway safety group is interested in studying the speed (km/hrs) of a car at a check point.

(2) The probability distribution of discrete r.v. X is as follows

$X = x$	1	2	3	4	5	6
$P(X = x)$	k	$2k$	$3k$	$4k$	$5k$	$6k$

(i) Determine the value of k. (ii) Find $P(X \le 4)$, $P(2 < X < 4)$, $P(X \ge 3)$.

(3) The following probability distribution of r.v. X

$X = x$	-3	-2	-1	0	1	2	3
$P(X = x)$	0·05	0·1	0·15	0·20	0·25	0·15	0.1

Find the probability that

(i) X is positive. (ii) X is non negative. (iii) X is odd. (iv) X is even.

(4) The p.m.f. of a r.v. X is given by $P(X = x) = \dfrac{({}^5C_x)}{2^5}$, for $x = 0, 1, 2, 3, 4, 5$ and $= 0$, otherwise.

Then show that $P(X \le 2) = P(X \ge 3)$.

(5) In the p.m.f. of r.v. X

x	1	2	3	4	5
$P(X)$	$\frac{1}{20}$	$\frac{3}{20}$	a	$2a$	$\frac{1}{20}$

Find a and obtain c.d.f. of X.

(6) A fair coin is tossed 4 times. Let X denotes the number of heads obtained write down the probability distribution of X. Also find the formula for p.m.f. of X.

(7) Find the probability distribution of the number of successes in two tosses of a die, where a success is defined as (i) number greater than 4 (ii) six appears on at least one die.

(8) A random variable X has the following probability distribution.

X	0	1	2	3	4	5	6	7
$P(X)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Determine (i) k (ii) $P(X > 6)$ (iii) $P(0 < X < 3)$

(9) The following is the c.d.f. of r.v. X

X	–3	–2	–1	0	1	2	3	4
$F(X)$	0·1	0·3	0·5	0·65	0·75	0·85	0·9	1

Find p.m.f. of X. (i) $P(-1 \le X \le 2)$ (ii) $P(X \le 3 / X > 0)$

(10) Find the expected value, variance and standard deviation of r.v. X whose p.m.f. are given below.

(i)

$X = x$	1	2	3
$P(X = x)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

(ii)

$X = x$	–1	0	1
$P(X = x)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

(iii)

$X = x$	1	2	3	. . .	n
$P(X = x)$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	. . .	$\frac{1}{n}$

(iv)

$X = x$	0	1	2	3	4	5
$P(X = x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

(11) A player tosses two coins he wins ₹ 10 if 2 heads appears , ₹ 5 if 1 head appears and ₹ 2 if no head appears. Find the expected winning amount and variance of winning amount.

(12) Let the p.m.f. of r.v. X be $P(x) = \frac{3-x}{10}$, for $x = -1, 0, 1, 2$ and $= 0$, otherwise

Calculate $E(X)$ and $Var(X)$.

(13) Suppose error involved in making a certain measurement is continuous r.v. X with p.d.f.

$f(x) = k(4 - x^2)$, for $-2 \le x \le 2$ and $= 0$ otherwise.

Compute (i) $P(X > 0)$ (ii) $P(-1 < x < 1)$ (iii) $P(X < 0.5$ or $X > 0.5)$

(14) The p.d.f. of r.v. X is given by $f(x) = \frac{1}{2a}$, for $0 < x < 2a$ and $= 0$, otherwise.

Show that $P\left(X < \frac{a}{2}\right) = P\left(X > \frac{3a}{2}\right)$.

(15) The p.d.f. of r.v. of X is given by $f(x) = \frac{k}{\sqrt{x}}$, for $0 < x < 4$ and $= 0$, otherwise.

Determine k . Determine c.d.f. of X and hence $P(X \le 2)$ and $P(X \le 1)$.