8 RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

8.1 Random Numbers In every day life, we base many of our decisions on random outcomes, i.e. chance occurrence. Captains of two teams toss a coin to decide as to which team will play first. Friends asking each other to offer a cup of tea or a cold drink write their names on slips of paper, fold them, mix them up and draw a slip. A person whose name is drawn in this way is required to serve the tea/drink. Lotteries are drawn by spinning a wheel. A scientific way of obtaining random outcomes is by the use of *random numbers*.

Random numbers are the numbers obtained by some random process (manual or mechanical). These numbers are assumed to be randomly and *uniformly* (equally) distributed. The basic random numbers are the ten one-digit numbers 0, 1, 2, ..., 9. Each of these numbers has an equal chance (1/10) of being selected. These numbers are combined into two-digit, three-digit, four-digit, ... numbers according to use. A two-digit random number table is given in Table 8.1 for easy reference.

Two-digit random numbers (00, 01, 02, ..., 09, 10, 11, ..., 99) are 100 in number and each of them has an equal probability (1/100) of being selected. Three-digit random numbers (000, 001, 002, ..., 010, 011, ..., 099, 100, 101, ..., 999) are 1000 in number and each of them has an equal probability (1/1000) of being selected. Four-digit numbers (0000, 0001, 0002, ..., 0100, 0101, ..., 0199, 1000, 1001, ..., 9999) are 10000 in number and each of them has an equal probability (1/10000) of being selected.

8.1.1 Generation of Random Numbers Random numbers can be generated manually as well as mechanically. Random numbers can be generated manually by drawing cards from playing cards or numbered cards, rotating or spinning numbered wheels etc. To generate random numbers, we can use the playing cards and exclude the pictured cards (Jack, Queen and King). After well-shuffling the pack, we draw a card and note the number drawn each time. For a card bearing number 10, we can write '0'. The invention of computers has very much simplified the generation of random numbers. Most of the electronic calculators have a key labelled *Random*.

The earlier attempts to generate random numbers were made by Karl Pearson in 1925 by studying census reports and by W.S. Gosset by shuffling and drawing 5000 cards. The widely used random numbers are:

1. Random sampling numbers by Tipper (1927). He generated 41,600 digits by drawing numbered cards from a bag.
2. Random sampling numbers by Kendall and Smith (1930). They generated 100,000 digits by rotating a disc.
3. Random numbers by Fisher and Yates (1938-39). They generated 15,000 digits based on entries from logarithmic table.
4. The largest and most ambitious generation of random numbers is by RAND Corporation (1955) based on the output of electronic pulsc generator contained in the book "A Million Random Numbers". This is a standard source.

8.1.2 Applications of Random Numbers The most common application of random numbers is for selection of samples. Random numbers are used for selection of random samples. Random numbers are also used wherever random selection is needed.

Suppose we want to select 10 students from a class of seventy five students. We shall use a two-digit random number table. To enter a table, we should close our eyes and place our finger or a ball-point on the table. Suppose the number below our finger or ball-point is 25. We should start from the fifth digit of the second row and go on noting the two-digit numbers obtained (ignoring numbers above 75). Suppose the first number is 37, second number is 84 (ignore it), third number is 42, fourth number is 63, and so on. In this manner we should choose 10 numbers.

A detailed procedure for selection of random samples will be explained in the chapter on *Sampling* in Chapter 11 (Book Part II)

8.2 Concept of a Random Variable We have already explained that an experiment is a process which generates raw data. Experiments in which outcomes vary from trial to trial are called random experiments. A variable whose values are determined by the outcomes of a random experiment is called a *random variable*. For example, 'throwing a die' is a random experiment and its outcome, i.e. the occurrence of 1, 2, 3, 4, 5 or 6 is a random variable. A random variable is also called a *chance variable* or a *stochastic variable*. We use capital letters X, Y or Z to denote a random variable and lower case letters x, y or z to denote its values.

Example 8.1 Suppose that a coin is tossed twice (or two coins are tossed simultaneously) so that the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads which can come up. With each sample point we can associate a number for X as shown in the table below. For example, in the case of HH (i.e. two heads) $X = 2$ while for TH (1 head) $X = 1$. It follows that X is a random variable.

Sample Point	HH	HT	TH	TT
X	2	1	1	0

8.2.1 Discrete and Continuous Random Variables A random variable which takes on only a finite number of values or a sequence of whole numbers is called a *discrete random variable*. For example, if we roll a die, the number of spots on the die is a discrete random variable. The number of accidents occurring on the G.T. Road during a month is a discrete random variable which could assume any of the possible values 0, 1, 2, ... A random variable which takes on an infinite number of values on a continuous scale in a given interval is called a *continuous random variable*. For example, the distance a shell will travel after being fired from a gun is a continuous random variable. A continuous random variable may assume any one of the values in the interval (a,b). It is important to remember that continuous random variables usually represent measurements data, e.g. heights, weights, temperatures, distances, lifetimes, etc., whereas discrete random variables represent count or enumeration data, e.g. number of books on a shelf, number of cars crossing Ravi bridge on a certain day, number of defective items in a lot, etc.

8.3 Discrete Probability Distributions Let a discrete random variable X assume values $x_1, x_2, \ldots, x_n$ with respective probabilities $P(x_1)$, $P(x_2)$, …, $P(x_n)$. An arrangement of all possible values of a random variable alongwith their respective probabilities is called a *probability distribution* or a *probability function*. Since the random variable takes a discrete set of values, it is also called a *discrete probability distribution*. A discrete probability distribution may take the form of a table, a graph or a mathematical equation. Whatever its form the discrete probability distribution must possess the following two properties:

1. $0 \leq P(x_i) \leq 1$, which means that probability lies between 0 and 1.
2. $\sum_i P(x_i) = 1$, which means that sum of probabilities is one.

Example 8.2 Suppose that a coin is tossed. Let 0 denote the occurrence of *head* and 1 denote the occurrence of *tail*. Thus the random variable X assumes the values 0 and 1. Since the coin is fair, the probability of head is 1/2 and that of tail is also 1/2. The probability function of X is thus given in the following table.

x	0	1
$P(x)$	1/2	1/2

Example 8.3 Suppose that a fair die is rolled. Here the random variable X assumes the value 1, 2, 3, 4, 5 or 6 (number of spots on the face of the die) each with probability 1/6. Thus the probability distribution of X is given in the following table.

x	1	2	3	4	5	6
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Example 8.4 If two coins are tossed and 'X' denotes the number of heads, construct probability distribution of X. (B.I.S.E., 2012)

Sol. When two coins are tossed the sample space will be $S = \{HH, HT, TH, TT\}$

Assuming that the coin is pair we have

$$P(HH) = \frac{1}{4},\ P(HT) = \frac{1}{4}$$

$$P(TH) \quad \text{and} \quad P(TT) = \frac{1}{4}$$

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

Probability function of X is given in the following table.

x	$P(x)$
0	1/4
1	1/2
2	1/4

Example 8.5 A bag contains two white and three black balls. Obtain a probability distribution of the number of white balls if two balls are selected.

(B.I.S.E., Multan 2000)

Solution Let the random variable X denote the number of white ball Then X assumes the values 0, 1 and 2.

$$P(X=0) = P(\text{no white ball}) = \frac{{}^2C_0\,{}^3C_2}{{}^5C_2} = \frac{1\times 3}{10} = \frac{3}{10}$$

$$P(X=1) = P(\text{one white ball}) = \frac{{}^2C_1\,{}^3C_1}{{}^5C_2} = \frac{2\times 3}{10} = \frac{6}{10}$$

$$P(X=2) = P(\text{two white balls}) = \frac{{}^2C_2\,{}^3C_0}{{}^5C_2} = \frac{1\times 1}{10} = \frac{1}{10}$$

The probability distribution of X is given in the following table.

x	0	1	2
$P(x)$	3/10	6/10	1/10

This probability distribution can be represented by the following probability function:

$$P(X=x) = \frac{{}^2C_x\,{}^3C_{2-x}}{{}^5C_2}, \qquad x = 0, 1, 2.$$

We have represented the probability function in both tabular form and as a mathematical equation. In Fig. 8.1(a), the function of Example 8.5 is displayed in its graphical form. Discrete data are often pictured by this type of *bar diagram*. The numbers corresponding to the values of the x's are located on the horizontal axis. Probabilities associated with the values of x are marked on the vertical axis. Above each possible x_i a vertical bar is erected; the height of this bar represents $P(x_i)$.

Instead of bar diagram, we more frequently construct rectangles, as in Fig. 8.1(b). Here the rectangles are constructed so that their bases of equal width are centred at each x value and their heights are equal to the corresponding probabilities given by $P(x_i)$. The bases are constructed so as to leave no space between the rectangles. Fig. 8.1(b) is called *probability histogram* or simply *histogram*.

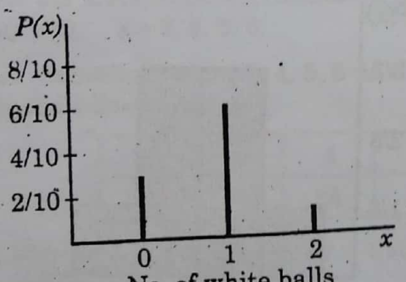


Fig. 8.1(a) Bar Diagram

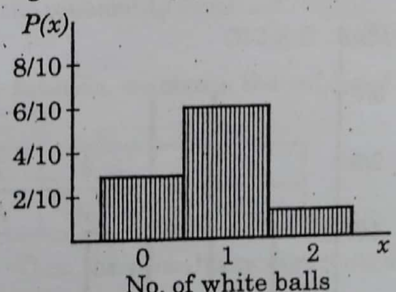


Fig. 8.1(b) Histogram

Example 8.6 Obtain a probability distribution for the number of heads when three coins are tossed. Also present the probability distribution by a graph.

Solution If H and T respectively denote the occurrence of head and tail, then the sample space for the experiment of tossing three coins is

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

If X denotes the number of heads when three coins are tossed, then the elements (sample points) of the sample space and the values of the random variable X are shown in the following table.

Sample Point	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X	3	2	2	2	1	1	1	0

The random variable X takes the value 0, 1, 2 or 3. The sample space consists of 8 sample points. There is only one sample point which corresponds to $X = 0$. Thus $P(X = 0) = 1/8$. There are three sample points which correspond to $X = 1$. Thus $P(X = 1) = 3/8$. Similarly, $P(X = 2) = 3/8$ and $P(X = 3) = 1/8$. Thus the probability distribution of the random variable X is given in the following table.

x	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

Alternative Solution Let X denote the number of heads. Then X assumes the value 0, 1, 2 or 3. Here $n = 3$, $p = 1/2$, $q = 1 - p = 1/2$. Applying the formula $P(X = x) = {}^nC_x p^x q^{n-x}$, we get

Binomial Distribution

$$P(X = 0) = {}^3C_0\,(1/2)^0\,(1/2)^3 = 1/8$$

$$P(X = 1) = {}^3C_1\,(1/2)\,(1/2)^2 = 3/8$$

$$P(X = 2) = {}^3C_2\,(1/2)^2\,(1/2) = 3/8$$

$$P(X = 3) = {}^3C_3\,(1/2)^3\,(1/2)^0 = 1/8$$

Thus we obtain the probability distribution of X as given in the above table. Figures 8.2(a) and 8.2(b) give respectively the bar diagram and the histogram for the above probability distribution.

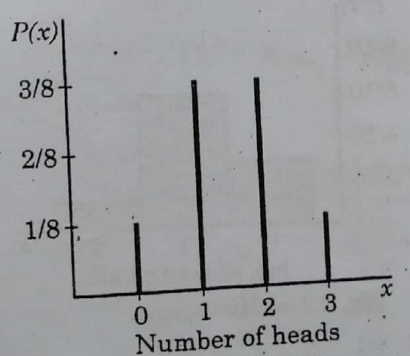


Fig 8.2(a) Bar Diagram

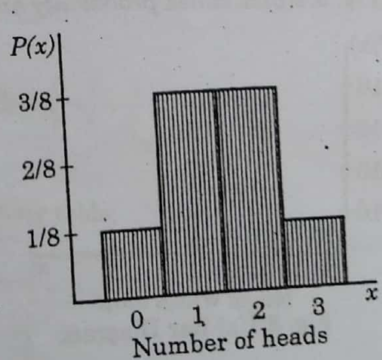


Fig. 8.2(b) Histogram

Example 8.7 Obtain a probability distribution of the sum of spots when a pair of dice is rolled.

Solution Let X denote a random variable whose values x are the possible sums of spots. Two dice can fall in $6 \times 6 = 36$ ways, each with probability 1/36. The sample points in the sample space and the corresponding values of X (sum of spots) are shown in Table 8.1.

The random variable X takes the values from 2 to 12. There are 36 sample points in the sample space. There is one sample point corresponding to $X = 2$. Thus $P(X = 2) = 1/36$. There are two sample points corresponding to $X = 3$. Thus $P(X = 3) = 2/36$. Similarly $P(X = 4) = 3/36, \ldots, P(X = 12) = 1/36$.

Table 8.1

Outcome of first die	Outcome of second die					
	1	2	3	4	5	6
1	(1,1)=2	(1,2)=3	(1,3)=4	(1,4)=5	(1,5)=6	(1,6)=7
2	(2,1)=3	(2,2)=4	(2,3)=5	(2,4)=6	(2,5)=7	(2,6)=8
3	(3,1)=4	(3,2)=5	(3,3)=6	(3,4)=7	(3,5)=8	(3,6)=9
4	(4,1)=5	(4,2)=6	(4,3)=7	(4,4)=8	(4,5)=9	(4,6)=10
5	(5,1)=6	(5,2)=7	(5,3)=8	(5,4)=9	(5,5)=10	(5,6)=11
6	(6,1)=7	(6,2)=8	(6,3)=9	(6,4)=10	(6,5)=11	(6,6)=12

The probability distribution of the random variable X is given in the following table.

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

As mentioned before, it is more convenient to show the probability function by means of a formula for $P(x)$. The mathematical equation for the probability function of X in this case is

$$P(X = x) = \frac{6 - |x - 7|}{36}, \qquad x = 2, 3, \ldots, 12.$$

Example 8.8 Determine the constant k in the probability function: (B.I.S.E., Lahore 2013)
$P(x) = k(x - 2), \quad x = 3, 4, 5, 6.$

Solution Substituting $x = 3, 4, 5, 6$ in the function, we obtain the values of $P(x)$ as shown in the following table.

x	3	4	5	6	
$P(x)$	k	$2k$	$3k$	$4k$	$\Sigma P(x) = 10k$

Now $\Sigma P(x) = 1$ or $10k = 1$ or $k = 1/10$. Thus the probability distribution of X is given in the following table.

x	3	4	5	6
$P(x)$	1/10	2/10	3/10	4/10

Example 8.9(a) An urn contains 5 white and 3 black balls. Two balls are drawn at random without replacement. If X denotes the number of white balls, then find the probability distribution of X. (B.I.S.E., Lahore 2015)

Solution The random variable X takes the value 0, 1 or 2.

$$P(X=0) = P(\text{no white ball}) = \frac{{}^5C_0\,{}^3C_2}{{}^8C_2} = \frac{1\times 3}{28} = \frac{3}{28}$$

$$P(X=1) = P(\text{one white ball}) = \frac{{}^5C_1\,{}^3C_1}{{}^8C_2} = \frac{5\times 3}{28} = \frac{15}{28}$$

$$P(X=2) = P(\text{two white balls}) = \frac{{}^5C_2\,{}^3C_0}{{}^8C_2} = \frac{10\times 1}{28} = \frac{5}{14}$$

The probability distribution of X is given in the following table.

x	0	1	2
$P(x)$	3/28	15/28	5/14

Example 8.9(b) In Example 8.9(a), if the balls are drawn with replacement, find the probability distribution of X.

Solution The random variable X takes the value 0, 1, or 2.

$$P(X=0) = P(\text{no white ball}) = P(\text{both balls black})$$

$$= P(\text{first ball black}) \cdot P(\text{second ball black}) = \frac{3}{8}\cdot\frac{3}{8} = \frac{9}{64}.$$

$$P(X=1) = P(\text{1 white ball})$$

$$= P(\text{first ball white and second ball black}) + P(\text{first ball black and second ball white})$$

$$= \frac{5}{8}\cdot\frac{3}{8} + \frac{3}{8}\cdot\frac{5}{8} = \frac{15}{64} + \frac{15}{64} = \frac{15}{32}.$$

$$P(X=2) = P(\text{both white balls}) = P(\text{first ball white}) \cdot P(\text{second ball white})$$

$$= \frac{5}{8}\cdot\frac{5}{8} = \frac{25}{64}.$$

The probability distribution of X is given in the following table.

x	0	1	2
$P(x)$	9/64	15/32	25/64

Example 8.10 Let X be a random variable giving the number of aces in a random draw of 4 cards from an ordinary deck of 52 cards. Construct a table showing the probability distribution of X.

Solution The random variable X takes the value 0, 1, 2, 3 or 4.

$$P(X=0)=P(\text{no ace})=\frac{{}^{4}C_0\,{}^{48}C_4}{{}^{52}C_4}=\frac{1\times 194580}{270725}=\frac{194580}{270725}$$

$$P(X=1)=P(1\text{ ace})=\frac{{}^{4}C_1\,{}^{48}C_3}{{}^{52}C_4}=\frac{4\times 17296}{270725}=\frac{69184}{270725}$$

$$P(X=2)=P(2\text{ aces})=\frac{{}^{4}C_2\,{}^{48}C_2}{{}^{52}C_4}=\frac{6\times 1128}{270725}=\frac{6768}{270725}$$

$$P(X=3)=P(3\text{ aces})=\frac{{}^{4}C_3\,{}^{48}C_1}{{}^{52}C_4}=\frac{4\times 48}{270725}=\frac{192}{270725}$$

$$P(X=4)=P(4\text{ aces})=\frac{{}^{4}C_4\,{}^{48}C_0}{{}^{52}C_4}=\frac{1\times 1}{270725}=\frac{1}{270725}$$

The probability distribution of X is given in the following table.

x	0	1	2	3	4
$P(x)$	$\frac{194580}{270725}$	$\frac{69184}{270725}$	$\frac{6768}{270725}$	$\frac{192}{270725}$	$\frac{1}{270725}$

Example 8.11 Find a formula for the probability distribution of the number of boys in families with three children assuming equal probabilities for boys and girls.

Solution There are $2^3 = 8$ points in the sample space representing equally likely outcomes, each with probability 1/8. To obtain the number of ways of getting 2 boys, for example, we need to consider the number of ways of partitioning 3 outcomes into two cells with two boys assigned to one cell and 1 girl to the other. This can be done in ${}^{3}C_2$ ways. In general, x boys and $3-x$ girls can occur in ${}^{3}C_x$ ways, where x assumes the value 0, 1, 2 or 3. Thus the probability function $P(X=x)=P(x)$ is given by

$$P(x)={}^{3}C_x\left(\frac{1}{8}\right),\quad x=0,1,2,3.$$

Putting $x = 0, 1, 2, 3$ in the above probability function, we obtain the probability distribution in tabular form as follows.

x	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

Example 8.12 Three marbles are drawn without replacement from an urn containing 4 red and 6 white marbles. Let X be a random variable that denotes the number of red marbles drawn. (a) Construct a table showing the probability distribution of X. (b) Find (i) $P(X=2)$ and (ii) $P(1 \le X \le 3)$, and interpret the results.

Solution (a) The random variable X takes the value 0, 1, 2 or 3.

$$P(X=0) = P(\text{no red marble}) = \frac{{}^4C_0\,{}^6C_3}{{}^{10}C_3} = \frac{1\times 20}{120} = \frac{1}{6}$$

$$P(X=1) = P(\text{1 red marble}) = \frac{{}^4C_1\,{}^6C_2}{{}^{10}C_3} = \frac{4\times 15}{120} = \frac{1}{2}$$

$$P(X=2) = P(\text{2 red marbles}) = \frac{{}^4C_2\,{}^6C_1}{{}^{10}C_2} = \frac{6\times 6}{120} = \frac{3}{10}$$

$$P(X=3) = P(\text{3 red marbles}) = \frac{{}^4C_3\,{}^6C_0}{{}^{10}C_3} = \frac{4\times 1}{120} = \frac{1}{30}$$

The probability distribution of X is given in the following table.

x	0	1	2	3
$P(x)$	1/6	1/2	3/10	1/30

(b)(i) $P(X=2) =$ 3/10.

This is the probability of drawing 2 red marbles.

(ii) $P(1 \le X \le 3) = P(X=1) + P(X=2) + P(X=3) = \frac{1}{2} + \frac{3}{10} + \frac{1}{30} = \frac{5}{6}$.

This is the probability of drawing 1, 2 or 3 red marble (i.e. the probability of drawing at least one red marble).

Note that a probability distribution is similar to a relative-frequency distribution with probabilities replacing the relative frequencies. Thus we can think of probability distributions as theoretical or ideal limiting forms of relative frequency distributions when the number of observations made is very large.

8.4 **Distribution Function for Discrete Random Variable:** In many problems, we may be interested to know the probability that the value of a random variable X is less than or equal to some number x. A function showing probabilities that a random variable X has a value less than or equal to x is called the *cumulative distribution function of X* or simply *distribution function of X*. Symbolically, the distribution function of X, denoted by $F(x)$, is given by $\boxed{F(x) = P(X \le x)}$. Thus $F(x)$ is the probability that X will assume a value less than or equal to x. It can be obtained from the probability function by noting that $F(x) = P(X \le x) = \sum_{u \le x} P(u)$ where the sum is taken over all values of u for which $u \le x$. Conversely, the probability function can be obtained from the distribution function.

The distribution function has the following properties:

1. $F(-\infty) = 0$ and $F(\infty) = 1$. This means that $F(x)$ is an increasing function and the range of $F(x)$ is the closed interval $[0, 1]$.
2. If $a < b$, then $F(a) < F(b)$ for any real numbers a and b.

For a discrete random variable, distribution function is obtained by cumulating probabilities just as we obtain cumulative frequency distribution from a frequency distribution by cumulating frequencies.

Example 8.13 The distribution function for the probability distribution of Example 8.5 is obtained in the following table.

x	$F(x)$
$x < 0$	0
$0 \le x < 1$	3/10
$1 \le x < 2$	9/10
$x \ge 2$	1

Example 8.14 The distribution function for the probability distribution of Example 8.6 is obtained in the following table.

x	$F(x)$
$x < 0$	0
$0 \le x < 1$	1/8
$1 \le x < 2$	4/8
$2 \le x < 3$	7/8
$x \ge 3$	1

Fig. 8.3 shows the graph of the distribution function for the above table.

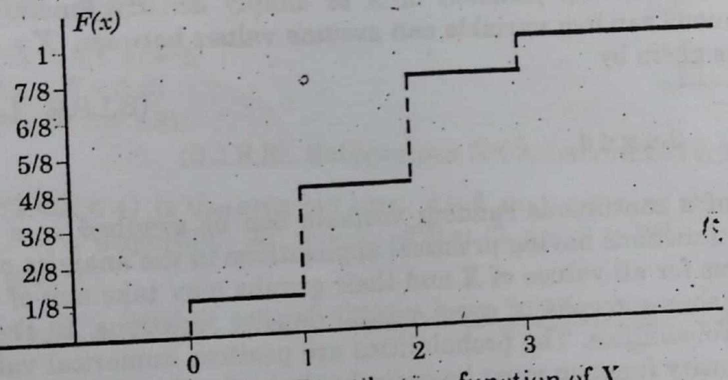


Fig. 8.3 Distribution function of X

Example 8.15 The distribution function for the probability distribution of Example 8.3 is obtained in the following table.

x	$x < 1$	$1 \le x < 2$	$2 \le x < 3$	$3 \le x < 4$	$4 \le x < 5$	$5 \le x < 6$	$x \ge 6$
$F(x)$	0	1/6	2/6	3/6	4/6	5/6	6/6

Example 8.16 The distribution function for the probability distribution of Example 8.7 is obtained in the following table.

x	$x<2$	$2\le x<3$	$3\le x<4$	$4\le x<5$	$5\le x<6$	$6\le x<7$
$F(x)$	0	1/36	3/36	6/36	10/36	15/36
x	$7\le x<8$	$8\le x<9$	$9\le x<10$	$10\le x<11$	$11\le x<12$	$x\ge 12$
$F(x)$	21/36	26/36	30/36	33/36	35/36	36/36

8.5 Continuous Probability Distributions We have seen that a random variable which can assume all possible values within a given interval is called a continuous random variable. There is an infinite number of values within a given interval of values. For example, there are an infinite number of weights between 69.5 kg and 70.5 kg. There is, however, very remote chance that a continuous random variable will assume a particular value. For example, it is very rare that the weight of a student selected at random will be exactly 70 kg. Thus the probability that a continuous random variable takes on a particular value is zero, i.e. $P(X=k)=0$.

We can, however, consider the probability of selecting a person who is at least 69 kg but not more than 70 kg, i.e. between 69 kg and 70 kg. Now we are dealing with an interval instead of a particular value of the random variable. In case of a continuous random variable, we compute probabilities for various intervals such as $P(a \le X \le b)$ or $P(W \ge c)$ and so on. It is important to note that when X is continuous $P(a \le X \le b) = P(a < X < b)$. This means that it does not matter whether we include an end point of the interval or not when the random variable is continuous.

Probability Density Function. The probability distribution of a continuous random variable cannot be presented in tabular form. It can be represented by means of a formula or a graph. The formula is necessarily in the form of a function of the numerical values of the continuous random variable X. In this case the function is called the *probability density function* of X or simply *density function* of X. For example, a continuous random variable can assume values between $X = 2$ and $X = 4$ and the function is given by

(B.I.S.E., Lahore 2014)

$$f(x) = \frac{x+1}{8}, \quad 2 \le x \le 4$$

A function of a continuous random variable can be graphed as a continuous curve. Most of the functions having practical applications in the analysis of statistical data are continuous for all values of X and their graphs may take any of the several forms. Figure 8.4 shows graphs of some typical density functions. In these graphs, areas represent probabilities. The probabilities are positive numerical values, which means that the density function must lie entirely above the X-axis.

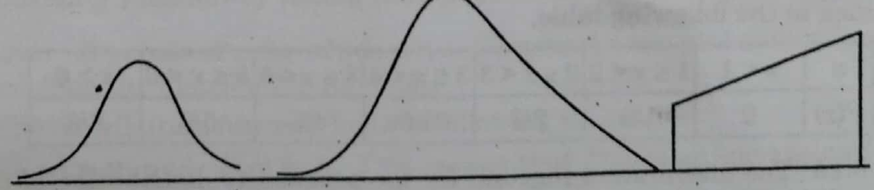


Fig. 8.4

A density function is constructed so that the area under its curve bounded by the X-axis is equal to 1. If a density function is represented by the curve as in Fig.8.5, then the probability that X assumes a value between a and b is the shaded area under the curve between the ordinates $x = a$ and $x = b$. The formal definition of a probability density functions is given below.

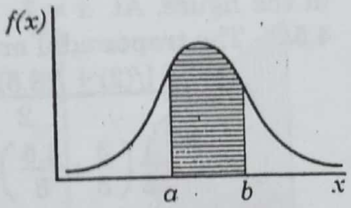


Fig. 8.5

A function $f(x)$ is called a *probability density function* for the continuous variable X if (i) it is non-negative and the total area under its curve and above the X-axis is equal to 1 (ii) the area under the curve between any two ordinates $x = a$ and $x = b$ gives the probability that X lies between a and b. Thus a probability density function has the following properties.

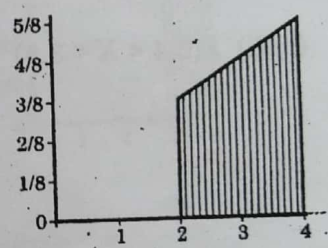
1. The function is non-negative, i.e. $f(x) \geq 0, \ a \leq x \leq b$.
2. The total area under the curve and above the X-axis is equal to 1, i.e.

$$\int_{-\infty}^{\infty} f(x)\, dx = 1.$$

In dealing with continuous distributions, we replace the summation notation used for discrete distributions by the integration notation.

Example 8.17 A continuous random variable X which assumes the values between $X = 2$ and $X = 4$ has the density function given by

$$f(x) = (x + 1)/8, \quad 2 \leq x \leq 4$$



(a) Show that $P(2 < X < 4) = 1$.
(b) Find (i) $P(2.4 < X < 3.5)$
(ii) $P(X < 3.5)$ (iii) $P(X = 1.5)$.

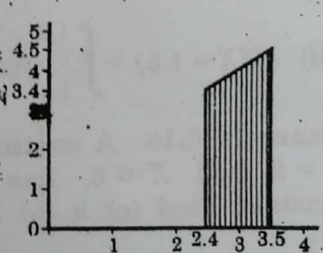
(B.I.S.E., Bahawalpur 2003; Lahore 2014; Multan 2018)

Solution (a) $P(2 < X < 4)$ is the area between $x = 2$ and $x = 4$ as shown shaded in the figure, which is a trapezoid. At $x = 2$, $f(2) = (2 + 1)/8 = 3/8$ and at $x = 4$, $f(4) = (4 + 1)/8 = 5/8$.

$$\text{Trapezoidal area} = \frac{(\text{sum of parallel sides}) \times \text{base}}{2} = \frac{[f(2) + f(4)]\,(4 - 2)}{2}$$

$$= \frac{1}{2}\left(\frac{3}{8} + \frac{5}{8}\right)(2) = 1.$$

(b) (i) The required probability is the area between $x = 2.4$ and $x = 3.5$ as shown shaded in the figure. At $x = 2.4$, $f(2.4) = (2.4 + 1)/8 = 3.4/8$ and at $x = 3.5$, $f(3.5) = (3.5 + 1)/8 = 4.5/8$. The trapezoidal area is



$$= \frac{[f(2.4) + f(3.5)] \times \text{base}}{2}$$

$$= \frac{1}{2}\left(\frac{3.4}{8} + \frac{4.5}{8}\right)(3.5 - 2.4) = \frac{1}{2}\left(\frac{7.9}{8}\right)(1.1) = \frac{8.69}{16} = 0.54.$$

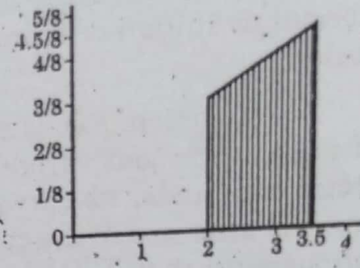
which is the required probability.

(ii) The required probability is the area between $x = 2$ and $x = 3.5$ as shown shaded in the figure. At $x = 2$, $f(2) = (2 + 1)/8 = 3/8$ and at $x = 3.5$, $f(3.5) = (3.5 + 1)/8 = 4.5/8$. The trapezoidal area is

$$\text{Area} = \frac{[f(2) + f(3.5)] \times \text{base}}{2}$$

$$= \frac{1}{2}\left(\frac{3}{8} + \frac{4.5}{8}\right)(3.5 - 2)$$

$$= \frac{1}{2}\left(\frac{7.5}{8}\right)(1.5) = 0.703,$$



which is the required probability.

(iii) $P(X = 1.5) = 0$, because the probability that a continuous random variable assumes a particular value is zero.

Alternative Solution

(a) $$P(2 < X < 4) = \int_2^4 f(x)\,dx = \frac{1}{8}\int_2^4 (1 + x)\,dx = \frac{1}{8}\left(x + \frac{x^2}{2}\right)\Big|_2^4$$

$$= \frac{1}{8}\left[\left(4 + \frac{16}{2}\right) - \left(2 + \frac{4}{2}\right)\right] = \frac{1}{8}(12 - 4) = 1.$$

(b) (i) $$P(2.4 < X < 3.5) = \int_{2.4}^{3.5} f(x)\,dx = \frac{1}{8}\int_{2.4}^{3.5} (1 + x)\,dx = \frac{1}{8}\left(x + \frac{x^2}{2}\right)\Big|_{2.4}^{3.5}$$

$$= \frac{1}{8}\left[\left(3.5 + \frac{(3.5)^2}{2}\right) - \left(2.4 + \frac{(2.4)^2}{2}\right)\right] = \frac{1}{8}\left(\frac{19.25}{2} - \frac{10.56}{2}\right)$$

$$= \frac{8.69}{16} = 0.54.$$

(ii) $$P(X < 3.5) = P(2 < X < 3.5) = \int_2^{3.5} f(x)\,dx = \frac{1}{8}\int_2^{3.5} (1 + x)\,dx$$

$$= \frac{1}{8}\left(x + \frac{x^2}{2}\right)\Big|_2^{3.5} = \frac{1}{8}\left[\left(3.5 + \frac{(3.5)^2}{2}\right) - \left(2 + \frac{4}{2}\right)\right]$$

$$= \frac{1}{8}\left(\frac{19.25}{2} - 4\right) = \frac{11.25}{16} = 0.703.$$

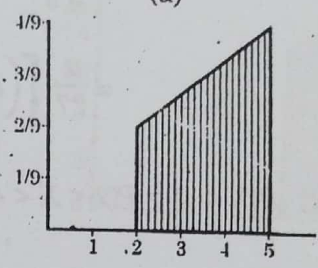
(iii) $$P(X = 1.5) = \int_{1.5}^{1.5} f(x)\,dx = 0$$

Example 8.18 A continuous random variable X, which assumes values between $X = 2$ and $X = 5$, has a density function given by $f(x) = k(1 + x)$, where k is a constant. Find (a) k (b) $P(3 < X < 4)$ and (c) $P(X < 4)$. (B.I.S.E., Lahore2018)

Solution (a) The density function of X is

$$f(x) = k(1 + x), \quad 2 < x < 5$$

(a)



The graph of $f(x) = k(1 + x)$ is a straight line as shown in the figure. We know that the total area under the line between $x = 2$ and $x = 5$ and above the X-axis must be 1. At $x = 2$, $f(2) = k(1 + 2) = 3k$ and at $x = 5$, $f(5) = k(1 + 5) = 6k$. Then we must choose k so that the trapezoidal area is 1. Hence we have

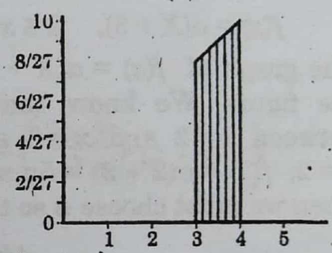
$$\text{Area} = \frac{[f(2) + f(5)]\ \text{base}}{2} = 1 \text{ or } \frac{1}{2}(3k + 6k)(5 - 2) = 1$$

or $27k/2 = 1$ or $k = 2/27$.

The density function of X is $f(x) = 2(1 + x)/27, \quad 2 < x < 5$

(b) The required probability is the area as shown shaded in the figure. Thus $f(3) = 2(1 + 3)/27 = 8/27$ and $f(4) = 2(1 + 4)/27 = 10/27$ are the ordinates (heights) at $x = 3$ and $x = 4$ respectively. The required trapezoidal area is

(b)



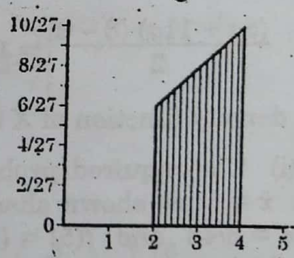
$$\text{Area} = \frac{[f(3) + f(4)] \times \text{base}}{2}$$

$$= \frac{1}{2}\left(\frac{8}{27} + \frac{10}{27}\right)(4 - 3) = \frac{1}{3},$$

which is the required probability.

(c) The required probability is the area between $x = 2$ and $x = 4$ as shown shaded in the figure. Thus $f(2) = 2(1 + 2)/27 = 2/9$ and $f(4) = 2(1 + 4)/27 = 10/27$ are the ordinates (heights) at $x = 2$ and $x = 4$ respectively. The required trapezoidal area is

©



$$\text{Area} = \frac{[f(2) + f(4)] \times \text{base}}{2} = \frac{1}{2}\left(\frac{2}{9} + \frac{10}{27}\right)(4 - 2) = \frac{16}{27},$$

which is the required probability.

Alternative Solution (a) The function $f(x)$ will be density function if (i) $f(x) \geq 0$ for every value of x, and (ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

For the second condition we must have $\int_2^5 f(x)\, dx = 1$ i.e.

$$\int_2^5 k(1 + x)\, dx = k\left(x + \frac{x^2}{2}\right)\Big|_2^5 = k\left[\left(5 + \frac{25}{2}\right) - \left(2 + \frac{4}{2}\right)\right]$$

$$= k\left(\frac{35}{2} - 4\right) = 1 \text{ or } \frac{27k}{2} = 1 \text{ or } k = \frac{2}{27}.$$

Thus the density function of X is $f(x) = 2(1 + x)/27, \; 2 < x < 5$

(b) $P(3 < X < 4) = \int_3^4 f(x)\,dx = \frac{2}{27}\int_3^4 (1+x)\,dx = \frac{2}{27}\left(x+\frac{x^2}{2}\right)\Big|_3^4$

$$= \frac{2}{27}\left[\left(4+\frac{16}{2}\right)-\left(3+\frac{9}{2}\right)\right] = \frac{2}{27}\left(12-\frac{15}{2}\right) = \frac{1}{3}.$$

(c) $P(X < 4) = P(2 < X < 4) = \int_2^4 f(x)\,dx = \frac{2}{27}\int_2^4 (1+x)\,dx$

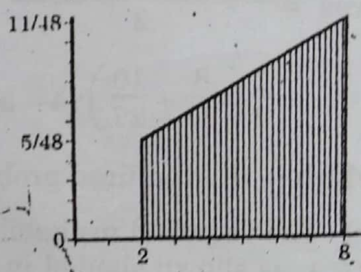
$$= \frac{2}{27}\left(x+\frac{x^2}{2}\right)\Big|_2^4 = \frac{2}{27}\left[\left(4+\frac{16}{2}\right)-\left(2+\frac{4}{2}\right)\right] = \frac{2}{27}(12-4) = \frac{16}{27}.$$

Example 8.19 A continuous random variable X that can assume values only between $X = 2$ and $X = 8$ has a density function given by $a(X + 3)$, where a is a constant
(a) Calculate a. (b) Find (i) $P(3 < X < 5)$ (ii) $P(X \geq 4)$ (iii) $P(|X - 5| < 0.5)$.
(B.I.S.E., Lahore 2003)

Solution (a) The density function of X is

$$f(x) = a(X+3), \quad 2 \leq x \leq 8$$

Example 8.19(a)

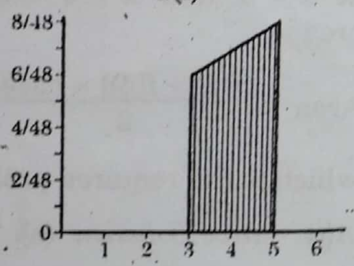


The graph of $f(x) = a(X + 3)$ is a straight line as shown in the figure. We know that the total area under the line between $x = 2$ and $x = 8$ and above the X-axis must be 1. At $x = 2$, $f(2) = a(2 + 3) = 5a$ and at $x = 8$, $f(8) = a(8 + 3) = 11a$. Then we must choose a so that the trapezoidal area is 1.

$$\text{Trapezoidal area} = \frac{[f(2)+f(8)] \times \text{base}}{2} = 1$$

or $$\frac{(5a+11a)(8-2)}{2} = 1 \text{ or } 48a = 1 \text{ or } a = 1/48.$$

b(i)



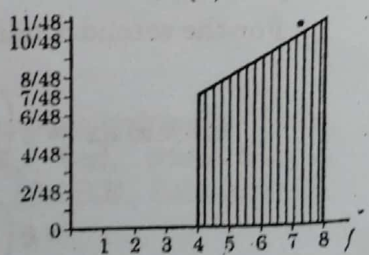
The density function of X is $f(x) = (x + 3)/48, \quad 2 \leq x \leq 8$

(b) (i) The required probability is the area between $x = 3$ and $x = 5$ as shown shaded in the figure. Thus $f(3) = (3 + 3)/48 = 6/48$ and $f(5) = (5 + 3)/48 = 8/48$ are the ordinates (heights) at $x = 3$ and $x = 5$ respectively. The required trapezoidal area is

$$\text{Area} = \frac{[f(3)+f(5)] \times \text{base}}{2} = \frac{1}{2}\left(\frac{6}{48}+\frac{8}{48}\right)(5-3) = \frac{7}{24},$$

which is the required probability.

b(ii)



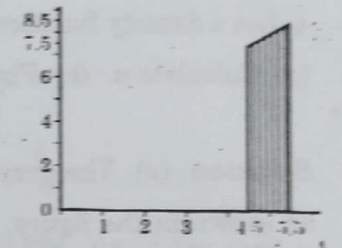
(ii) The required probability is the area between $x = 4$ and $x = 8$ as shown shaded in the figure. Thus $f(4) = (4 + 3)/48 = 7/48$ and $f(8) = (8 + 3)/48 = 11/48$ are the ordinates (heights) at $x = 3$ and $x = 8$ respectively. The required trapezoidal are is

$$\text{Area} = \frac{[f(4) + f(8)] \times \text{base}}{2} = \frac{1}{2}\left(\frac{7}{48} + \frac{11}{48}\right)(8-4) = \frac{3}{4},$$

which is the required probability.

(iii) $P(|X-5| < 0.5) = P(-0.5 < (X-5) < 0.5) = P(4.5 < X < 5.5)$; that is, the required probability is the area between $x = 4.5$ and $x = 5.5$ as shown shaded in the figure. Thus $f(4.5) =$ $(4.5 + 3)/48 = 7.5/48$ and $f(5.5) = (5.5 + 3)/48 = 8.5/48$ are the ordinates (heights) at $x = 4.5$ and $x = 5.5$ respectively. The required trapezoial area is

(iii)



$$\text{Area} = \frac{[f(4.5) + f(5.5)]}{2} = \frac{1}{2}\left(\frac{7.5}{48} + \frac{8.5}{48}\right)(5.5 - 4.5) = \frac{1}{6},$$

which is the required probability.

Alternative Solution (a) The function $f(x)$ will be a density function if (i) $f(x) \geq 0$ for every value of x; and (ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$. For the second condition, we must have $\int_{2}^{8} f(x)\, dx = 1$. Now

$$\int_{2}^{8} a(x+3)\, dx = a\left(\frac{x^2}{2} + 3x\right)\Big|_{2}^{8} = a\left[\left(\frac{8 \times 8}{2} + (3 \times 8)\right) - \left(\frac{2 \times 2}{2} + (3 \times 2)\right)\right]$$

$$= a(56 - 8) = 48a.$$

Thus $48a = 1$ or $a = 1/48$.

The density function of X is $f(x) = (x + 3)/48, \quad 2 \leq x \leq 8$

(b) (i) $$P(3 < X < 5) = \int_{3}^{5} f(x)\, dx = \frac{1}{48}\int_{3}^{5} (x+3)\, dx = \frac{1}{48}\left(\frac{x^2}{2} + 3x\right)\Big|_{3}^{5}$$

$$= \frac{1}{48}\left[\left(\frac{25}{2} + 15\right) - \left(\frac{9}{2} + 9\right)\right] = \frac{7}{24}$$

(ii) $$P(X \geq 4) = P(4 \leq X \leq 8) = \int_{4}^{8} f(x)\, dx = \frac{1}{48}\int_{4}^{8} (x+3)\, dx = \frac{1}{48}\left(\frac{x^2}{2} + 3x\right)\Big|_{4}^{8}$$

$$= \frac{1}{48}\left[\left(\frac{64}{2} + 24\right) - \left(\frac{16}{2} + 12\right)\right] = \frac{3}{4}.$$

(iii) $$P(|X-5| < 0.5) = P(4.5 < X < 5.5) = \int_{4.5}^{5.5} f(x)\, dx$$

$$= \frac{1}{48}\int_{4.5}^{5.5} (x+3)\, dx = \frac{1}{48}\left(\frac{x^2}{2} + 3x\right)\Big|_{4.5}^{5.5}$$

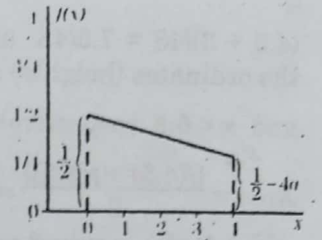
$$= \frac{1}{48}\left[\left(\frac{30.25}{2} + 16.5\right) - \left(\frac{20.25}{2} + 13.5\right)\right] = \frac{1}{48}\left(\frac{63.25}{2} - \frac{47.25}{2}\right) = \frac{1}{6}$$

Example 8.20 A continuous random variable X having values only between 0 and 4 has a density function given by $f(x) = \frac{1}{2} - ax$, where a is a constant.

(a) Calculate a. (b) Find $P(1 < X < 2)$.

(B.I.S.E., Multan 1999; Gujranwala 2003, 2014; Lahore 2011)

Solution (a) The graph of $f(x) = \frac{1}{2} - ax$ is a straight line as shown in the figure. We know that the total area under the line between $x = 0$ and $x = 4$ and above the X-axis must be 1. At $x = 0$, $f(0) = \frac{1}{2}$,

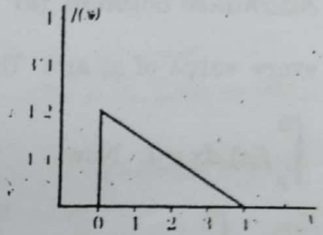


and at $x = 4$, $f(4) = \frac{1}{2} - 4a$. Thus we choose a so that the trapezoidal area is 1.

$$\text{Area} = \frac{[f(0) + f(4)] \times \text{base}}{2}$$

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

Now $2(1 - 4a) = 1$ or $a = 1/8$.

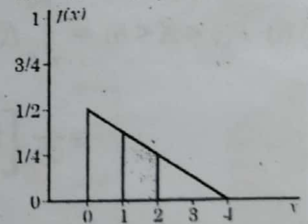


Thus $\left(\frac{1}{2} - 4a\right)$ is actually equal to zero, and so the correct graph is as shown in the figure.

$\frac{1}{2} - ax$

(b). The required probability is the area between $x = 1$ and $x = 2$ as shown shaded in the figure. From part (a), $f(x) = \frac{1}{2} - \frac{1}{8}x$,

thus $f(1) = \frac{3}{8}$ and $f(2) = \frac{1}{4}$ are the ordinates (heights) at $x = 1$ and $x = 2$ respectively. The required trapezoidal area is given by



$$\text{Area} = \frac{[f(1) + f(2)] \times \text{base}}{2} = \frac{1}{2}\left(\frac{3}{8} + \frac{1}{4}\right)(2 - 1) = \frac{5}{16},$$

which is the required probability.

Alternative Solution (a) The function $f(x)$ will be a density function if (i) $f(x) \geq 0$ for every value of x; and (ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

$$\int_0^4 \left(\frac{1}{2} - ax\right) dx = \left(\frac{x}{2} - \frac{ax^2}{2}\right)\Bigg|_0^4 = \left(\frac{4}{2} - \frac{16a}{2}\right) - 0 = 2 - 8a.$$

Thus $2 - 8a = 1$ or $a = 1/8$.

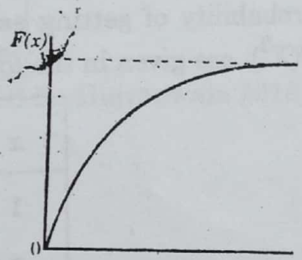
The density function of X is $f(x) = \frac{1}{2} - \frac{1}{8}x$, $0 < x < 4$.

(b) $P(1 < X < 2) = \int_1^2 f(x)\,dx = \int_1^2 \left(\frac{1}{2} - \frac{1}{8}x\right) dx = \left(\frac{x}{2} - \frac{x^2}{2(8)}\right)\Big|_1^2$

$$= \left[\left(\frac{2}{2} - \frac{4}{16}\right) - \left(\frac{1}{2} - \frac{1}{16}\right)\right] = 1 - \frac{1}{4} - \frac{1}{2} + \frac{1}{16} = \frac{5}{16}.$$

8.6 Distribution Function for Continuous Random Variable As in the case discrete random variable, the cumulative distribution function or the distribution function for the continuous random variable X is defined as

$$F(x) = P(X \le x) = P(-\infty \le X \le x) = \int_{-\infty}^{x} f(t)\,dt \;.$$



The distribution function is an increasing function which increases from 0 to 1 and is represented by a curve as shown in the figure.

Note that the derivative of an indefinite integral of a function is the function itself, i.e.

$$\frac{d}{dx} F(x) = \frac{d}{dx} \int_{-\infty}^{x} f(t)\,dt = f(x).$$

The distribution function can be used to find $P(a \le X \le b)$ as

$$P(a \le X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a).$$

8.7 Mathematical Expectation Let a random variable X assume the values $x_1, x_2, \ldots, x_n$ with respective probabilities $P(x_1), P(x_2), \ldots, P(x_n)$ such that the sum of the probabilities is equal to 1, i.e. $\Sigma P(x_i) = 1$. The *mathematical expectation* or *expected value* of X, denoted by $E(X)$, is defined as

$$E(X) = x_1 P(x_1) + x_2 P(x_2) + \cdots + x_n P(x_n) = \sum_{i=1}^{n} x_i P(x_i) = \Sigma x P(x) \qquad (8.1)$$

The expectation of X is usually called the *mean of* X and is denoted by μ. Whenever several random variables, such as $X, Y, \ldots$ are being studied together, the symbol μ should be subscripted to indicate the particular variable involved, that is, $E(X) = \mu_x$ and $E(Y) = \mu_y$. When only one variable is being studied the subscripts are often omitted.

Example 8.21 Find $E(X)$ and $E(X^2)$ for the probability function of X in Example 8.5.

Solution Calculations needed for finding $E(X)$ and $E(X^2)$ are given in following table.

x	$P(x)$	$x\,P(x)$	$x^2\,P(x)$
0	3/10	0	0
1	6/10	6/10	6/10
2	1/10	2/10	4/10
		$\Sigma xP(x) = 8/10$	$\Sigma x^2P(x) = 1$

$E(X) = \Sigma x\, P(x) = 8/10$ and $E(X^2) = \Sigma x^2\, P(x) = 10/10 = 1$.

Note that the expected value of a discrete random variable is not always an integer and not always one of the values that the random variable can assume. It represents that constant around which the values of the random variable oscillate. It is simply the *weighted mean* of the possible values, the weights being the probabilities of occurrence for these values.

Example 8.22 If X represents the spot on a die, find $E(X)$ and $E(X^2)$.

Solution A die has six faces having 1, 2, 3, 4, 5, 6 spots. If the die is fair, the probability of getting each spot is 1/6. Calculations needed for finding $E(X)$ and $E(X^2)$ are given in the following table.

x	$P(x)$	$x\,P(x)$	$x^2\,P(x)$
1	1/6	1/6	1/6
2	1/6	2/6	4/6
3	1/6	3/6	9/6
4	1/6	4/6	16/6
5	1/6	5/6	25/6
6	1/6	6/6	36/6
		$\Sigma xP(x) = 21/6$	$\Sigma x^2P(x) = 91/6$

$E(X) = 21/6 = 3.5$ and $E(X^2) = 91/6 = 15.17$.

Example 8.23 The probability distribution of a discrete random variable X is given by

$$f(x) = {}^3C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \qquad x = 0, 1, 2, 3$$

Find $E(X)$ and $E(X^2)$. (B.I.S.E., Sargodha 2003; Bahawalpur 2003; Faisalabad 2017)

Solution On substituting $x = 0, 1, 2, 3$ in the probability function we obtain the values of the function $f(x)$ as shown in the table below. Calculation of $E(X)$ and $E(X^2)$ is also outlined below.

x	0	1	2	3	
$f(x)$	27/64	27/64	9/64	1/64	$\Sigma f(x) = 64/64 = 1$
$xf(x)$	0	27/64	18/64	3/64	$\Sigma xf(x) = 48/64$
$x^2f(x)$	0	27/64	36/64	9/64	$\Sigma x^2f(x) = 72/64$

$E(X) = \Sigma x\, f(x) = \dfrac{48}{64} = \dfrac{3}{4}$ and $E(X^2) = \Sigma x^2\, f(x) = \dfrac{72}{64} = \dfrac{9}{8}$.

Example 8.23(b) From the following probability distribution find mean and variance.

X	P(X = X)	XP(x)	$X^2P(x)$
0	1/16	0	0
1	4/16	4/16	4/16
2	6/16	12/16	24/16
3	4/16	12/16	36/16
4	1/16	4/16	16/16
	$\Sigma P(x) = 1$	$\Sigma XP(x) = 2$	$\Sigma X^2P(x) = 5$

(B.I.S.E., Gujranwala 2018)

Solution. $\mu = E(X) \quad = \Sigma XP(x) = 2$

$$E(X^2) \quad = \Sigma X^2P(x) = 5$$

$$Var(X) = E(X^2) - [E(X)]^2 = 5 - (2)^2 = 5 - 4 = 1$$

Example 8.24 A bag contains 4 white and 3 red balls. A person draws two balls at random without replacement being promised Rs. 14 for each white ball and Rs. 7 for each red ball he draws. Find his expectation.

Solution Two balls drawn can be both white, one white and one red or both red. The person receives Rs. 28 if he draws both white balls, Rs. 21 if he draws one white and one red balls and Rs. 14 if he draws both red balls. The respective probabilities of the drawings are

$$P(\text{two white balls}) = \frac{{}^4C_2\,{}^3C_0}{{}^7C_2} = \frac{6 \times 1}{21} = \frac{2}{7}$$

$$P(\text{one white and one red balls}) = \frac{{}^4C_1\,{}^3C_1}{{}^7C_2} = \frac{4 \times 3}{21} = \frac{4}{7}$$

$$P(\text{two red balls}) = \frac{{}^3C_2\,{}^4C_0}{{}^7C_2} = \frac{3 \times 1}{21} = \frac{1}{7}$$

Hence the required expectation is

$$E(X) = x_1\,P(x_1) + x_2\,P(x_2) + x_3\,P(x_3) = 28\left(\frac{2}{7}\right) + 21\left(\frac{4}{7}\right) + 14\left(\frac{1}{7}\right) = 8 + 12 + 2$$
$$= 22.$$

Example 8.25 A bag contains 2 white and 3 black balls. Four persons A, B, C and D in the order named each draws one ball and does not replace it. The first to draw a white ball receives Rs. 10. Determine their expectations.

Solution Let $P(A_1)$ denote the probability that A draws a white ball on the first draw and $P(A_1')$ denote the probability that A does not draw a white ball on the first draw. Let $P(B_2)$ denote the probability that B draws a white ball on the second draw and $P(B_2')$ denote the probability that B does not draw a white ball on the second draw. Let $P(C_3)$, $P(C_3')$ and $P(D_4)$ denote respectively the probabilities that C draws a white ball on the third draw, C' does not draw a white ball on the third draw and D draws a white ball on the fourth draw. Then

$$P(A) = P(A_1) = \frac{2}{5}$$

$$P(B) = P(A_1') P(B_2) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right) = \frac{3}{10}$$

$$P(C) = P(A_1') P(B_2') P(C) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{2}{3}\right) = \frac{1}{5}$$

$$P(D) = P(A_1') P(B_2') P(C_3') P(D_4) = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)\left(\frac{2}{2}\right) = \frac{1}{10}$$

Hence their respective expectations are

$E(A) = 10\,(2/5) =$ Rs.4 $\quad E(B) = 10(3/10) =$ Rs.3

$E(C) = 10(1/5) =$ Rs.2 $\quad E(D) = 10(1/10) =$ Re.1

Example 8.26 A and B toss a coin alternately on the understanding that the first to throw a head will win Rs. 15. Find their expectations of money if A has the first throw.

Solution From Example 7.36, $P(A) = 2/3$ and $P(B) = 1/3$. Hence their expectations of money are

$E(A) = 15(2/3) =$ Rs.10 and $E(B) = 15(1/3) =$ Rs.5.

Example 8.27 Three men take turns in throwing a die for a prize of Rs.91 to be given to the one who first throws a '6'. Find their expectations.

Solution From Example 7.37, $P(A) = 36/91$, $P(B) = 30/91$ and $P(C) = 25/91$. Hence their expectations are

$E(A) = 91(36/91) =$ Rs. 36, $E(B) = 91(30/91) =$ Rs. 30,
$E(C) = 91(25/91) =$ Rs. 25.

8.7.1 Laws of Expectation

1. If X is a random variable and a and b constants, then

   $$E(aX + b) = a\,E(X) + b \tag{8.2}$$

   From this law, we have the following results.

   (i) If $a = 0$, then $E(b) = b$, i.e. the expectation or expected value of a constant is equal to the constant itself.

   (ii) If $b = 0$, then $E(aX) = a\,E(X)$, i.e. the expected value of the product of a constant and a random variable is equal to the product of the constant and the expected value of the random variable.

2. The expected value of the sum or difference of two random variables is equal to the sum or difference of the expected values of the individual random variables. Symbolically, if X and Y are random variables, then $E(X + Y) = E(X) + E(Y)$ and $E(X - Y) = E(X) - E(Y)$.

3. The expected value of the product of two independent random variables is equal to the product of their individual expected values. Symbolically, if X and Y are independent random variables, then $E(XY) = E(X)\,E(Y)$.

4. The expected value of the deviation of a random variable from its own expected value is equal to zero. Symbolically, if $E(X)$ is the expected value of X, then $E[X - E(X)] = 0$.

Example 8.28 For the probability distribution of Example 8.5, show that $E(5X+3) = 5E(X) + 3$.

Solution Necessary calculations are given in the following table.

x	$P(x)$	$5x+3$	$(5x+3)\,P(x)$
0	3/10	3	9/10
1	6/10	8	48/10
2	1/10	13	13/10
			$\Sigma(5x+3)\,P(x) = 70/10$

$E(5X+3) = \Sigma(5X+3)\,P(x) = 70/10 = 7.$

From Example 8.21, we have $E(X) = \frac{8}{10}$;

$5E(X) + 3 = 5\left(\frac{8}{10}\right) + 3 = 7.$

Thus $E(5X+3) = 5E(X) + 3.$

Example 8.29 A random variable X has the probability function:

x	-2	3	1
$f(x)$	1/3	1/2	1/6

Find (a) $E(X)$, $E(2X+5)$, $E(X^2)$, $E(2x+5)^2$. (b) Show that $E(2X+5) = 2E(X) + 5$.

Solution Calculations required for finding $E(X)$, $E(X^2)$, $E(2X+5)$ and $E(2X+5)^2$ are shown in the following table.

x	$P(x)$	$xP(x)$	$x^2P(x)$	$2x+5$	$(2x+5)P(x)$	$(2x+5)^2P(x)$
-2	1/3	$-2/3$	4/3	1	1/3	1/3
3	1/2	3/2	9/2	11	11/2	121/2
1	1/6	1/6	1/6	7	7/6	49/6
		$\Sigma xP(x)=6/6$	$\Sigma x^2P(x)=36/6$		$\Sigma(2x+5)P(x)=7$	$\Sigma(2x+5)^2P(x)=414/6$

(a) $E(X) = \Sigma x\,P(x) = 6/6 = 1$

$E(2X+5) = \Sigma(2x+5)\,P(x) = 42/6 = 7$

$E(X^2) = \Sigma x^2\,P(x) = 36/6 = 6$

$E(2X+5)^2 = \Sigma(2x+5)^2\,P(x) = 414/6 = 69$

(b) $2E(X) + 5 = 2(1) + 5 = 7$

Thus $E(2X+5) = 2E(X) + 5$.

Example 8.30 In summer season a dealer of desert room coolers earns 500 per day if the day is hot and Rs. 200 per day if it is fair and loses Rs. 60 per day if it is cloudy. Find his mathematical expectation if the probability of the day being hot is 0.40 and for being fair is 0.25. (B.I.S.E., Rawalpindi 2002)

Solution Let the random variable X denote the amount the dealer can earn. Let X_h denote the amount the dealer can earn on a hot day, X_f denote the amount the dealer can earn on a fair day and X_c denote the amount the dealer can lose on a cloudy day, so that $X = X_h + X_f - X_c$. Now

$$P(\text{hot}) = P(X_h) = 0.40,\ P(\text{fair}) = P(X_f) = 0.25$$

$$P(\text{cloudy}) = P(X_c) = 1 - (0.40 + 0.25) = 0.35. \text{ Thus}$$

$$E(\text{hot}) = E(X_h) = 500(0.40) = 200,$$

$$E(\text{fair}) = E(X_f) = 200(0.25) = 50 \text{ and } E(\text{cloudy}) = E(X_c) = 60(0.35) = 21. \text{ Hence}$$

$$E(X) = E(X_h) + E(X_f) - E(X_c) = 200 + 50 - 21 = \text{Rs. } 229.$$

8.8 The Variance and Standard Deviation The variance of a random variable X, denoted by $\text{Var}(X)$ or σ^2, is defined as $\text{Var}(X) = E[X - E(X)]^2 = E[(X - \mu)^2]$

The symbols $\text{Var}(X)$ and σ^2 are often used interchangeably to denote the variance of a random variable. If more than one variable is being considered, then σ (sigma) should be subscripted as $\text{Var}(X) = \sigma_x^2$ and $\text{Var}(Y) = \sigma_y^2$; otherwise the subscripts are omitted.

The variance is a non-negative number and its units of measurement are square units of the original scale. The positive square root of the variance is called the *standard deviation* and is given by

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{E[(X - \mu)^2]}.$$

If X is a discrete random variable with probability function $P(x)$, then the variance is given by

$$\sigma^2 = E[(X - \mu)^2] = \sum_{i=1}^{n} (x_i - \mu)^2 P(x_i) = \Sigma(x - \mu)^2 P(x) \qquad (8.3)$$

In the special case of (8.3) where the probabilities are equal i.e.

$P(x_1) = P(x_2) = \cdots = P(x_n) = \frac{1}{n}$, we have

$$\sigma^2 = \frac{1}{n}[(x_1 - \mu)^2 + \cdots + (x_n - \mu)^2] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

which is called the variance of a set of n numbers $x_1, \ldots, x_n$.

For calculating variances the following identity is often useful.

$$\left.\begin{aligned} \text{Var}(X) &= E[X - E(X)]^2 = E(X^2) - (E(X))^2 \\ &= \sum_{i=1}^{n} x_i^2 P(x_i) - \mu^2 = \Sigma x^2 P(x) - \mu^2 \end{aligned}\right\} \qquad (8.4)$$

Example 8.31(a) Find the variance and standard deviation for the probability function of X in Example 8.5.

Solution Calculations required for determining $\text{Var}(X)$ and σ are shown in the following table.

Scanned with CamScanner

x	$P(x)$	$x\,P(x)$	$x^2P(x)$	$x-\mu$	$(x-\mu)^2$	$(x-\mu)^2\,P(x)$
0	3/10	0	0	−0.8	0.64	0.192
1	6/10	6/10	6/10	0.2	0.04	0.024
2	1/10	2/10	4/10	1.2	1.44	0.144
		$\Sigma xP(x)=8/10$	$\Sigma x^2P(x)=10/10=1$			$\Sigma(2x-\mu)^2P(x)=0.36$

$\mu = \Sigma x\,P(x) = 8/10 = 0.8$ and $\sigma^2 = \Sigma(x-\mu)^2\,P(x) = 0.36$.

Alternately, $\sigma^2 = \Sigma x^2\,P(x) - \mu^2 = 1-(0.8)^2 = 1-0.64 = 0.36$,

and $\sigma = \sqrt{0.36} = 0.6$.

Example 8.31(b) $E(x^2) = 400$, var $(X) = 144$. Find $E(X)$

Sol. Variance $= E(X^2) - [E(X)]^2$

$$144 = 400 - [E(X)]^2 \quad \Rightarrow \quad [E(X)]^2 = 400-144 \quad \Rightarrow \quad [E(X)]^2 = 256$$

Taking square root on both sides

$$\sqrt{[E(X)]^2} = \sqrt{256} \quad \Rightarrow \quad E(X) = 16$$

8.8.1 Properties of Variance For any random variable X, Var $(X) \geq 0$. The following laws are often useful in computing variance. If X is a random variable and both a and b are constants, then

(i) $\text{Var}(a) = 0$

(ii) $\text{Var}(X \pm a) = \text{Var}(X)$

(iii) $\text{Var}(aX) = a^2\,\text{Var}(X)$

(iv) $\text{Var}(a+bX) = b^2\,\text{Var}(X)$

The properties of standard deviation follow by taking the square root of the results for variance.

Example 8.32 For the probability distribution in Example 8.29, find the variance and standard deviation of X. Also find $\text{Var}(2X+5)$ and $\text{S.D.}(2X+5)$ and discuss their relations with $\text{Var}(X)$ and $\text{S.D.}(X)$.

Solution From Example 8.29, $\mu = E(X) = 1$, $E(X^2) = 6$, $E(2X+5) = 7$ and $E(2X+5)^2 = 69$.

$$\text{Var}(X) = E(X^2) - \mu^2 = 6-(1)^2 = 5$$

and $\text{S.D.}(X) = \sqrt{5} = 2.236$.

$$\text{Var}(2X+5) = E(2X+5)^2 - [E(2X+5)]^2 = 69-(7)^2 = 20$$

and $\text{S.D.}(2X+5) = 4.472$.

Relations $\text{Var}(2X+5) = (2)^2\,\text{Var}(X) = 4(5) = 20$.

$$\text{S.D.}(2X+5) = 2[\text{S.D.}(X)] = 2(2.236) = 4.472.$$

Scanned with CamScanner

Example 8.33 A continuous random variable X has a density function given by $A(x+5)$. Find A and $P(3 < X < 4)$, if range of variable is $3 < x < 4$.

(B.I.S.E., Rawalpindi, 2010)

Solution $f(x) = A(x+5) \qquad 3 < x < 4$

$$\int_3^4 f(x) = dx = 1$$

$$\int_3^4 A(x+5)\,dx = 1$$

$$A \left| \frac{x^2}{2} + 5x \right|_3^4 = 1$$

$$A\left[\frac{(4)^2}{2} + 5(4) - \frac{(3)^2}{2} + (5 \times 3)\right] = 1$$

$$A\left[(8+20) - \left(\frac{9}{2} + 15\right)\right] = 1$$

$$A\left(28 - \frac{39}{2}\right) = 1$$

$$A\,(17/2) = 1$$

$$A = \frac{2}{17}$$

The density function of X is $f(x)$

$$f(x) = \frac{2\,(X+5)}{17} \qquad 3 < X < 4$$

(i) $$P(3 < X < 4) = \frac{2}{17}\int_3^4 (x+5)\,dx = \frac{2}{17}\frac{x^2}{2} + 5x \Big|_3^4 = \frac{2}{17}\left[\left(\frac{16}{2} + 20\right) - \left(\frac{9}{2} + 15\right)\right]$$

$$= \frac{2}{17}\left[(8+20) - \left(\frac{39}{2}\right)\right] = \frac{2}{17}\left[28 - \frac{39}{2}\right] = \frac{2}{17}\left[\frac{56-39}{2}\right]$$

$$= \frac{2}{17}\left[\frac{17}{2}\right] = 1$$

Example 8.34 The probability density function of a random variable with range. $f(x) = a(3x+2) \quad 2 \le x \le 4$. Find (i) a (ii) $P(0 \le x \le 2.7)$. (B.I.S.E., Gujranwala, 2010)

Solution $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

$$\int_2^4 a(3x+2) = 1$$

$$a \left| \frac{3x^2}{2} + 2x \right|_2^4 = 1$$

$$a \left[\left(\frac{3(4)^2}{2} + 2(4)\right) - \left(\frac{3(2)^2}{2} + 2(2)\right)\right] = 1$$

$$a \left[\left(\frac{48}{2} + 8\right) - \left(\frac{12}{2} + 4\right)\right] = 1$$

$$a\ [(24 + 8) - (6 + 4)] = 1$$

$$a\ [32 - 10] = 1 \qquad a(22) = 1$$

$$\boldsymbol{a = 1/22} \qquad f(x) = 3x + \frac{2}{22} \quad 2 \le x \le 4$$

(ii) $P(0 \le x\ 2.7) = \int_2^{2.7} f(x)\,dx = \int_2^{2.7} \frac{(3x+2)}{22}\,dx = \frac{1}{22} \left| \frac{3x^2}{2} + 2x \right|_2^{2.7}$

$$= \frac{1}{22} \left[\left(\frac{3(2.7)^2}{2} + 2(2.7)\right) - \left(\frac{3(2)^2}{2} - 2(2)\right)\right]$$

$$= \frac{1}{22}\ [(10.935 + 5.4) - (6 - 2)] = \frac{1}{22}\ [(16.335 - 2)] = \frac{14.335}{22} = 0.6516$$