26.7 RANDOM VARIABLE

If a real variable X be associated with the outcome of a random experiment, then since the values which X takes depend on chance, it is called a *random variable* or a *stochastic variable* or simply a *variate*. For instance, if a random experiment E consists of tossing a pair of dice, the sum X of the two numbers which turn up have the value 2, 3, 4, ..., 12 depending on chance. Then X is the random variable. It is a function whose values are real numbers and depend on chance.

If in a random experiment, the event corresponding to a number a occurs, then the corresponding random variable X is said to assume the value a and the probability of the event is denoted by $P(X = a)$. Similarly the probability of the event X assuming any value in the interval $a < X < b$ is denoted by $P(a < X < b)$. The probability of the event $X \leq c$ is written as $P(X \leq c)$.

If a random variable takes a finite set of values, it is called a *discrete variate*. On the other hand, if it assumes an infinite number of uncountable values, it is called a *continuous variate*.

26.8 (1) DISCRETE PROBABILITY DISTRIBUTION

Suppose a discrete variate X is the outcome of some experiment. If the probability that X takes the values x_i is p_i, then

$$P(X = x_i) = p_i \text{ or } p(x_i) \text{ for } i = 1, 2, \ldots$$

where (*i*) $p(x_i) \geq 0$ for all values of i, (*ii*) $\Sigma p(x_i) = 1$

The set of values x_i with their probabilities p_i constitute a **discrete probability distribution** *of the discrete variate X.*

For example, the discrete probability distribution for X, the sum of the numbers which turn on tossing a pair of dice is given by the following table :

$X = x_i$	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

[$\because$ There are $6 \times 6 = 36$ equally likely outcomes and therefore, each has the probability 1/36. We have $X = 2$ for one outcome, *i.e.* (1, 1) ; $X = 3$ for two outcomes (1, 2) and (2, 1); $X = 4$ for three outcomes (1, 3), (2, 2) and (3, 1) and so on.]

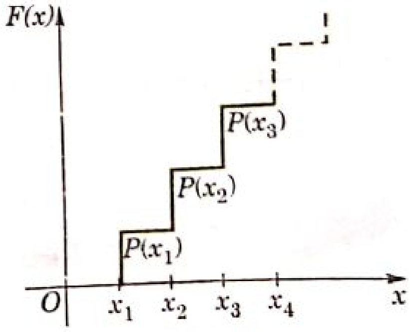

Fig. 26.2

(2) Distribution function. *The distribution function $F(x)$ of the discrete variate X is defined by*

$$F(x) = P(X \leq x) = \sum_{i=1}^{x} p(x_i) \text{ where } x \text{ is any integer. The graph of } F(x) \text{ will be}$$

stair step form (Fig. 26.2). The distribution function is also sometimes called *cumulative distribution function*.

Example 26.28. *A die is tossed thrice. A success is 'getting 1 or 6' on a toss. Find the mean and variance of the number of successes.* (V.T.U., 2011 S ; Rohtak, 2004)

Solution. Probability of a success $= \frac{2}{6} = \frac{1}{3}$, Probability of failures $= 1 - \frac{1}{3} = \frac{2}{3}$.

$\therefore$ prob. of no success = Prob. of all 3 failures $= \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}$

Probability of one successes and 2 failures $= 3c_1 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$

Probability of two successes and one failure $= 3c_2 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$

Probability of three successes $= \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$

Now				
	$x_i = 0$	1	2	3
	$p_i = 8/27$	4/9	2/9	1/27

$\therefore$ mean $\mu = \Sigma p_i x_i = 0 + \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = 1.$

Also $\Sigma p_i x_i^2 = 0 + \frac{4}{9} + \frac{8}{9} + \frac{9}{27} = \frac{5}{3}$

$\therefore$ variance $\sigma^2 = \Sigma p_i x_i^2 - \mu^2 = \frac{5}{3} - 1 = \frac{2}{3}$.

Example 26.29. *The probability density function of a variate X is*

X :	0	1	2	3	4	5	6
$p(X)$:	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

(i) Find $P(X < 4)$, $P(X \geq 5)$, $P(3 < X \leq 6)$. (V.T.U., 2013)

(ii) What will be the minimum value of k so that $P(X \leq 2) > 3$.

Solution. (*i*) If X is a random variable, then

$$\sum_{i=0}^{6} p(x_i) = 1 \text{ i.e., } k + 3k + 5k + 7k + 9k + 11k + 13k = 1 \text{ or } k = 1/49.$$

$\therefore$ $P(X < 4) = k + 3k + 5k + 7k = 16k = 16/49.$

$P(X \geq 5) = 11k + 13k = 24k = 24/49.$

$P(3 < X \leq 6) = 9k + 11k + 13k = 33k = 33/49.$

(*ii*) $P(X \leq 2) = k + 3k + 5k = 9k > 0.3$ or $k > 1/30$

Thus minimum value of $k = 1/30$.

Example 26.30. *A random variable X has the following probability function :*

x :	0	1	2	3	4	5	6	7
$p(x)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

(i) Find the value of the k *(ii) Evaluate* $P(X < 6)$, $P(X \geq 6)$ (V.T.U., 2014)

(iii) $P(0 < X < 5)$. (Anna, 2012 ; W.B.T.U., 2005 ; J.N.T.U., 2003)

Solution. (*i*) If X is a random variable, then

$$\sum_{i=0}^{7} p(x_i) = 1, \text{ i.e., } 0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

i.e., $7k^2 + 9k - 1 = 0$ *i.e.* $(10 - k)(k + 1) = 0$ *i.e.,* $k = \frac{1}{10}$

(*ii*) $P(X < 6) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$= 0 + k + 2k + 2k + 3k + k^2 = 8k + k^2 = \frac{8}{10} + \frac{1}{100} = \frac{81}{100}$

$$P(X \geq 6) = P(X = 6) + P(X = 7) = 2k^2 + 7k^2 + k = \frac{9}{100} + \frac{1}{10} = \frac{19}{100}$$

(ii) $P(0 < X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= k + 2k + 2k + 3k = 8k = \frac{8}{10} = \frac{4}{5}.$$

26.9 (1) CONTINUOUS PROBABILITY DISTRIBUTION

When a variate X takes every value in an interval, it gives rise to *continuous distribution* of X. The distributions defined by the variates like heights or weights are continuous distributions.

A major conceptual difference, however, exists between discrete and continuous probabilities. When thinking in discrete terms, the probability associated with an event is meaningful. With continuous events, however, where the number of events is infinitely large, the probability that a specific event will occur is practically zero. For this reason, continuous probability statements must be worded somewhat differently from discrete ones. Instead of finding the probability that x equals some value, we find the probability of x falling in a small interval.

Thus the probability distribution of a continuous variate x is defined by a function $f(x)$ such that the probability of the variate x falling in the small interval $x - \frac{1}{2}dx$ to $x + \frac{1}{2}dx$ is $f(x)\,dx$. Symbolically it can be expressed as $P\left(x - \frac{1}{2}dx \leq x \leq x + \frac{1}{2}dx\right) = f(x)\,dx$. Then $f(x)$ is called the *probability density function* and the continuous curve $y = f(x)$ is called the *probability curve.*

The range of the variable may be finite or infinite. But even when the range is finite, it is convenient to consider it as infinite by supposing the density function to be zero outside the given range. Thus if $f(x) = \phi(x)$ be the density function denoted for the variate x in the interval (a, b), then it can be written as

$$\begin{aligned} f(x) &= 0, && x < a \\ &= \phi(x), && a \leq x \leq b \\ &= 0, && x > b. \end{aligned}$$

The density function $f(x)$ is always positive and $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (*i.e.*, the total area under the probability curve and the x-axis is unity which corresponds to the requirements that the total probability of happening of an event is unity).

(2) Distribution function

If $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx,$

then $F(x)$ *is defined as the* **cumulative distribution function** *or simply the* **distribution function** *of the continuous variate* X. It is the probability that the value of the variate X will be $\leq x$. The graph of $F(x)$ in this case is as shown in Fig. 26.3(b).

The distribution function $F(x)$ has the following properties :

(i) $F'(x) = f(x) \geq 0$, so that $F(x)$ is a non-decreasing function.

(i) $F(-\infty) = 0$; (iii) $F(\infty) = 1$

(iv) $P(a \leq x \leq b) = \int_{a}^{b} f(x)\,dx = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx = F(b) - F(a).$

Example 26.31. (i) *Is the function defined as follows a density function?*

$$\begin{aligned} f(x) &= e^{-x}, && x \geq 0 \\ &= 0, && x < 0, \end{aligned}$$

(ii) *If so, determine the probability that the variate having this density will fall in the interval (1, 2) ?*

(iii) *Also find the cumulative probability function F(2) ?*

Solution. (i) $f(x)$ is clearly ≥ 0 for every x in (1, 2) and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{\infty} e^{-x}\,dx = 1$$

Hence the function $f(x)$ satisfies the requirements for a density function.

(*ii*) Required probability $= P(1 \le x \le 2) = \int_1^2 e^{-x}\,dx = e^{-1} - e^{-2} = 0.368 - 0.135 = 0.233.$

This probability is equal to the shaded area in Fig. 26.3 (*a*).

(*iii*) Cumulative probability function $F(2)$

$$\int_{-\infty}^{2} f(x)\,dx = \int_{-\infty}^{0} 0 \cdot dx + \int_0^2 e^{-x}\,dx = 1 - e^{-2} = 1 - 0.135 = 0.865$$

which is shown in Fig. 26.3 (*b*).

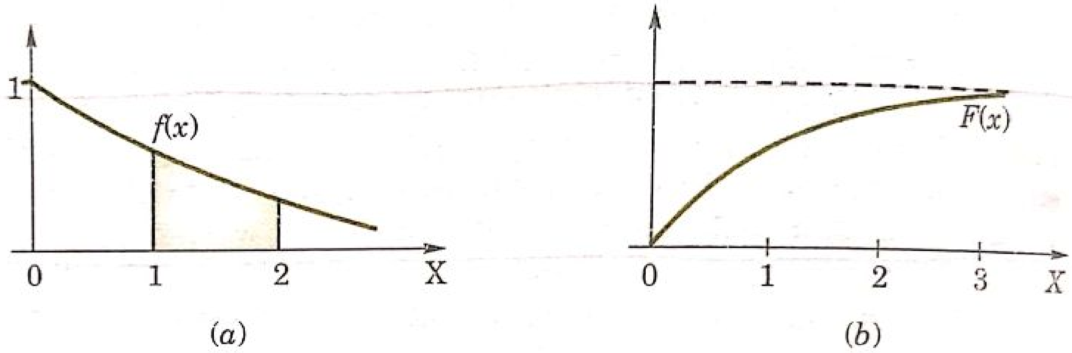

Fig. 26.3