UNIT 5 RANDOM VARIABLES

Structure

5.1 INTRODUCTION

In the previous units, we have studied the assignment and computation of probabilities of events in detail. In those units, we were interested in knowing the occurrence of outcomes. In the present unit, we will be interested in the numbers associated with such outcomes of the random experiments. Such an interest leads to study the concept of random variable.

In this unit, we will introduce the concept of random variable, discrete and continuous random variables in Sec. 5.2 and their probability functions in Secs. 5.3 and 5.4.

Objectives

A study of this unit would enable you to:

- define a random variable, discrete and continuous random variables;
- specify the probability mass function, i.e. probability distribution of discrete random variable;
- specify the probability density function, i.e. probability function of continuous random variable; and
- define the distribution function.

5.2 RANDOM VARIABLE

Study related to performing the random experiments and computation of probabilities for events (subsets of sample space) have been made in detail in the first four units of this course. In many experiments, we may be interested in a numerical characteristic associated with outcomes of a random experiment. Like the outcome, the value of such a numerical characteristic cannot be predicted in advance.

For example, suppose a die is tossed twice and we are interested in number of times an odd number appears. Let X be the number of appearances of odd number. If a die is thrown twice, an odd number may appear ‘0’ times (i.e. we may have even number both the times) or once (i.e. we may have odd number in one throw and even number in the other throw) or twice (i.e. we may have odd number both the times). Here, X can take the values 0, 1, 2 and is a variable quantity behaving randomly and hence we may call it as 'random variable'. Also notice that its values are real and are defined on the sample space

{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 1), (5, 2), (5, 3), (5, 4), (5, 4), (5, 6),
(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

i.e.

$$X = \begin{cases} 0, \text{ if the outcome is } (2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2), \\ \qquad (6,4),(6,6) \\ 1, \text{ if the outcome is } (1,2),(1,4),(1,6),(2,1),(2,3),(2,5),(3,2), \\ \qquad (3,4),(3,6),(4,1),(4,3),(4,5),(5,2),(5,4), \\ \qquad (5,6),(6,1),(6,3),(6,5) \\ 2, \text{ if the outcome is } (1,1),(1,3),(1,5),(3,1),(3,3),(3,5),(5,1), \\ \qquad (5,3),(5,5) \end{cases}$$

So, $P[X = 0] = \frac{9}{36} = \frac{1}{4}, P[X = 1] = \frac{18}{36} = \frac{1}{2}, P[X = 2] = \frac{9}{36} = \frac{1}{4}$,

and $P[X = 0] + P[X = 1] + P[X = 2] = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$

Observe that, a probability can be assigned to the event that X assumes a particular value. It can also be observed that the sum of the probabilities corresponding to different values of X is one.

So, a random variable can be defined as below:

Definition: A random variable is a real-valued function whose domain is a set of possible outcomes of a random experiment and range is a sub-set of the set of real numbers and has the following properties:

i) Each particular value of the random variable can be assigned some probability

ii) Uniting all the probabilities associated with all the different values of the random variable gives the value 1(unity).

Remark 1: We shall denote random variables by capital letters like X, Y, Z, etc. and write r.v. for random variable.

5.3 DISCRETE RANDOM VARIABLE AND PROBABILITY MASS FUNCTION

Discrete Random Variable

A random variable is said to be discrete if it has either a finite or a countable number of values. Countable number of values means the values which can be arranged in a sequence, i.e. the values which have one-to-one correspondence with the set of natural numbers, i.e. on the basis of three-four successive known terms, we can catch a rule and hence can write the subsequent terms. For example suppose X is a random variable taking the values say 2, 5, 8, 11, … then we can write the fifth, sixth, … values, because the values have one-to-one correspondence with the set of natural numbers and have the general term as $3n - 1$ i.e. on taking n = 1, 2, 3, 4, 5, … we have 2, 5, 8, 11, 14,…. So, X in this example is a discrete random variable. The number of students present each day in a class during an academic session is an example of discrete random variable as the number cannot take a fractional value.

Probability Mass Function

Let X be a r.v. which takes the values x_1, x_2, ... and let $P\left[X = x_i\right] = p(x_i)$. This function $p(x_i)$, i =1,2, … defined for the values x_1, x_2, … assumed by X is called probability mass function of X satisfying $p(x_i) \geq 0$ and $\sum_i p\left(x_i\right) = 1$.

The set $\left\{\left(x_1, p\left(x_1\right)\right), \left(x_2, p\left(x_2\right)\right), ...\right\}$ specifies the probability distribution of a discrete r.v. X. Probability distribution of r.v. X can also be exhibited in the following manner:

X	x_1	x_2	x_3 …
$p(x)$	$p(x_1)$	$p(x_2)$	$p(x_3)$…

Now, let us take up some examples concerning probability mass function:

Example 1: State, giving reasons, which of the following are not probability distributions:

(i)

X	0	1
$p(x)$	$\frac{1}{2}$	$\frac{3}{4}$

(ii)

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

(iii)

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

(iv)

X	0	1	2	3
p(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{8}$

Solution:

(i) Here $p(x_i) \geq 0$, i = 1, 2; but

$$\sum_{i=1}^{2} p(x_i) = p(x_1) + p(x_2) = p(0) + p(1) = \frac{1}{2} + \frac{3}{4} = \frac{5}{4} > 1.$$

So, the given distribution is not a probability distribution as $\sum_{i=1}^{2} p(x_i)$ is greater than 1.

(ii) It is not probability distribution as $p(x_2) = p(1) = -\frac{1}{2}$ i.e. negative

(iii) Here, $p(x_i) \geq 0$, i = 1, 2, 3

and $\sum_{i=1}^{3} p(x_i) = p(x_1) + p(x_2) + p(x_3) = p(0) + p(1) + p(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$

$\therefore$ The given distribution is probability distribution.

(iv) Here, $p(x_i) \geq 0$, i = 1, 2, 3, 4; but

$$\sum_{i=1}^{4} p(x_i) = p(x_1) + p(x_2) + p(x_3) + p(x_4)$$

$$= p(0) + p(1) + p(2) + p(3) = \frac{1}{8} + \frac{3}{8} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} < 1.$$

$\therefore$ The given distribution is not probability distribution.

Example 2: For the following probability distribution of a discrete r.v. X, find

i) the constant c,

ii) $P[X \leq 3]$ and

iii) $P[1 < X < 4]$.

X	0	1	2	3	4	5
$p(x)$	0	c	c	2c	3c	c

Solution:

i) As the given distribution is probability distribution,

$$\therefore \quad \sum_i p(x_i) = 1$$

$$\Rightarrow 0 + c + c + 2c + 3c + c = 1 \Rightarrow 8c = 1 \Rightarrow c = \frac{1}{8}$$

ii) $P[X \le 3] = P[X = 3] + P[X = 2] + P[X = 1] + P[X = 0]$

$$= 2c + c + c + 0 = 4c = 4 \times \frac{1}{8} = \frac{1}{2}.$$

iii) $P[1 < X < 4] = P[X = 2] + P[X = 3] = c + 2c = 3c = 3 \times \frac{1}{8} = \frac{3}{8}.$

Example 3: Find the probability distribution of the number of heads when three fair coins are tossed simultaneously.

Solution: Let X be the number of heads in the toss of three fair coins.

As the random variable, "the number of heads" in a toss of three coins may be 0 or 1 or 2 or 3 associated with the sample space

{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT},

$\therefore$ X can take the values 0, 1, 2, 3, with

$P[X = 0] = P[TTT] = \frac{1}{8}$

$P[X = 1] = P[HTT, THT, TTH] = \frac{3}{8}$

$P[X = 2] = P[HHT, HTH, THH] = \frac{3}{8}$

$P[X = 3] = P[HHH] = \frac{1}{8}.$

$\therefore$ Probability distribution of X, i.e. the number of heads when three coins are tossed simultaneously is

X	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

which is the required probability distribution.

Example 4: A r.v. X assumes the values $-2, -1, 0, 1, 2$ such that

$P[X = -2] = P[X = -1] = P[X = 1] = P[X = 2],$

$P[X < 0] = P[X = 0] = P[X > 0].$

Obtain the probability mass function of X.

Solution: As $P[X < 0] = P[X = 0] = P[X > 0]$

$\therefore P[X = -1] + P[X = -2] = P[X = 0] = P[X = 1] + P[X = 2]$

$\Rightarrow p + p = P[X = 0] = p + p$

[Letting $P[X = 1] = P[X = 2] = P[X = -1] = P[X = -2] = p$]

$\Rightarrow P[X = 0] = 2p.$

Now, as $P[X < 0] + P[X = 0] + P[X > 0] = 1$,

$\therefore P[X = -1] + P[X = -2] + P[X = 0] + P[X = 1] + P[X = 2] = 1$

$\Rightarrow p + p + 2p + p + p = 1$

$\Rightarrow 6p = 1 \Rightarrow p = \frac{1}{6}$

$\therefore P[X = 0] = 2p = 2 \times \frac{1}{6} = \frac{2}{6}$,

$P[X = -1] = P[X = -2] = P[X = 1] = P[X = 2] = p = \frac{1}{6}$.

Hence, the probability distribution of X is given by

X	-2	-1	0	1	2
p(x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Now, here are some exercises for you.

E1) 2 bad articles are mixed with 5 good ones. Find the probability distribution of the number of bad articles, if 2 articles are drawn at random.

E2) Given the probability distribution:

X	0	1	2	3
p(x)	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{10}$

Let $Y = X^2 + 2X$. Find the probability distribution of Y.

E3) An urn contains 3 white and 4 red balls. 3 balls are drawn one by one with replacement. Find the probability distribution of the number of red balls.

Let us define and explain a continuous random variable and its probability function in the next section.

5.4 CONTINUOUS RANDOM VARIABLE AND PROBABILITY DENSITY FUNCTION

In Sec. 5.3 of this unit, we have defined the discrete random variable as a random variable having countable number of values, i.e. whose values can be arranged in a sequence. But, if a random variable is such that its values cannot be arranged in a sequence, it is called continuous random variable. Temperature of a city at various points of time during a day is an example of continuous random variable as the temperature takes uncountable values, i.e. it can take fractional values also. So, a random variable is said to be continuous if it can take all possible real (i.e. integer as well as fractional) values between two certain limits. For example, let us denote the variable, "Difference between the rainfall (in cm) of a city and that of another city on every rainy day in a rainy reason", by X, then X here is a continuous random variable as it can take any real value between two certain limits. It can be noticed that for a continuous random variable, the chance of occurrence of a particular value of the variable is very small, so instead of specifying the probability of taking a particular value by the variable, we specify the probability of its lying within an interval. For example, chance that an athlete will finish a race in say exactly 10 seconds is very-very small, i.e. almost zero as it is very rare to finish the race in a fixed time. Here, the probability is specified for an interval, i.e. we may be interested in finding as to what is the probability of finishing the race by the athlete in an interval of say 10 to 12 seconds.

So, continuous random variable is represented by different representation known as **probability density function** unlike the discrete random variable which is represented by probability mass function.

Probability Density Function

Let $f(x)$ be a continuous function of x. Suppose the shaded region ABCD shown in the following figure represents the area bounded by $y = f(x)$, x-axis and the ordinates at the points x and $x + \delta x$, where δx is the length of the interval $(x, x + \delta x)$.

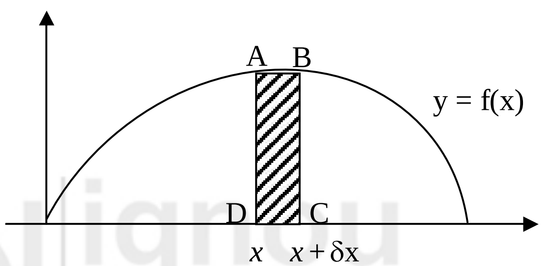


Fig. 5.1

Now, if δx is very-very small, then the curve AB will act as a line and hence the shaded region will be a rectangle whose area will be AD × DC i.e. $f(x)\,\delta x$ [$\because$ AD = the value of y at x i.e. $f(x)$, DC = length δx of the interval $(x, x+\delta x)$]

Also, this area = probability that X lies in the interval $(x, x+\delta x)$

$$= P[x \le X \le x + \delta x]$$

Hence,

$$P[x \le X \le x + \delta x] = f(x)\delta x$$

$$\Rightarrow \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x), \text{ where } \delta x \text{ is very-very small}$$

$$\Rightarrow \lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x).$$

$f(x)$, so defined, is called probability density function.

Probability density function has the same properties as that of probability mass function. So, $f(x) \ge 0$ and sum of the probabilities of all possible values that the random variable can take, has to be 1. But, here, as X is a continuous random variable, the summation is made possible through 'integration' and hence

$$\int_R f(x)\,dx = 1,$$

where integral has been taken over the entire range R of values of X.

Remark 2

i) Summation and integration have the same meanings but in mathematics there is still difference between the two and that is that the former is used in case of discrete values, i.e. countable values and the latter is used in continuous case.

ii) An essential property of a continuous random variable is that there is zero probability that it takes any specified numerical value, but the probability that it takes a value in specified intervals is non-zero and is calculable as a definite integral of the probability density function of the random variable and hence the probability that a continuous r.v. X will lie between two values a and b is given by

$$P[a < X < b] = \int_a^b f(x)\,dx.$$

Example 5: A continuous random variable X has the probability density function:

$f(x) = Ax^3$, $0 \le x \le 1$.

Determine

i) A

ii) $P[0.2 < X < 0.5]$

iii) $P[X > \frac{3}{4} \text{ given } X > \frac{1}{2}]$

Solution:

(i) As $f(x)$ is probability density function,

$$\therefore \int_R f(x)\,dx = 1$$

$$\Rightarrow \int_0^1 f(x)\,dx = 1 \Rightarrow \int_0^1 Ax^3 dx = 1$$

$$\Rightarrow A\left[\frac{x^4}{4}\right]_0^1 = 1 \Rightarrow A\left(\frac{1}{4} - 0\right) = 1 \Rightarrow A = 4$$

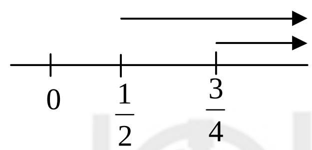
(ii) $P[0.2 < X < 0.5] = \int_{0.2}^{0.5} f(x)\,dx = \int_{0.2}^{0.5} Ax^3 dx = 4\left[\frac{x^4}{4}\right]_{0.2}^{0.5} = [(0.5)^4 - (0.2)^4]$

$$= 0.0625 - 0.0016 = 0.0609$$

(iii) $P[X > \frac{3}{4} \text{ given } X > \frac{1}{2}] = P[X > \frac{3}{4} \mid X > \frac{1}{2}]$

$$= \frac{P\left[X > \frac{3}{4} \cap X > \frac{1}{2}\right]}{P\left[X > \frac{1}{2}\right]} \qquad \left[\because P(A|B) = \frac{P(A \cap B)}{P(B)}\right]$$

$$= \frac{P\left[X > \frac{3}{4}\right]}{P\left[X > \frac{1}{2}\right]} \qquad \left[\begin{array}{l}\because \text{the common portion for} \\ X > \frac{3}{4} \text{ and } X > \frac{1}{2} \text{ is } X > \frac{3}{4}\end{array}\right]$$



Now, $P\left[X > \frac{3}{4}\right] = \int_{3/4}^{1} f(x)\,dx$

$$\left[\begin{array}{l}\because \text{ lower limit is } \frac{3}{4} \text{ and upper limit} \\ \text{is given in the problem which is } 1\end{array}\right]$$

$$= \int_{3/4}^{1} 4x^3 dx = 4\left[\frac{x^4}{4}\right]_{3/4}^{1} = (1)^4 - \left(\frac{3}{4}\right)^4 = 1 - \frac{81}{256} = \frac{175}{256}, \text{ and}$$

$$P\left[X > \frac{1}{2}\right] = \int_{1/2}^{1} f(x) = [x^4]_{1/2}^{1} = 1 - \frac{1}{16} = \frac{15}{16}.$$

$$\therefore \text{ the required probability } = \frac{P\left[X > \frac{3}{4}\right]}{P\left[X > \frac{1}{2}\right]} = \frac{175}{256} \times \frac{16}{15} = \frac{35}{16 \times 3} = \frac{35}{48}.$$

Example 6: The p.d.f. of the different weights of a "1 litre pure ghee pack" of a company is given by:

$$f(x) = \begin{cases} 200(x-1) & \text{for } 1 \le x \le 1.1 \\ 0, & \text{otherwise} \end{cases}$$

Examine whether the given p.d.f. is a valid one. If yes, find the probability that the weight of any pack will lie between 1.01 and 1.02.

Solution: For $1 \le x \le 1.1$, we have $f(x) \ge 0$, and

$$\int_1^{1.1} f(x)\,dx = \int_1^{1.1} 200(x-1)\,dx = 200\left[\frac{x^2}{2}-x\right]_1^{1.1} = 200\left[\left\{\frac{(1.1)^2}{2}-1.1\right\}-\left\{\frac{1}{2}-1\right\}\right]$$

$$= 200\left[\left(\frac{1.21-2.2}{2}\right)-\left(\frac{1-2}{2}\right)\right] = 200\left[-\frac{0.99}{2}+\frac{1}{2}\right] = 200\frac{(0.01)}{2} = 1.$$

$\therefore$ f(x) is p.d.f.

$$\text{Now, } P[1.01 < X < 1.02] = \int_{1.01}^{1.02} 200(x-1)\,dx = 200\left[\frac{x^2}{2}-x\right]_{1.01}^{1.02}$$

$$= 200\left[\left\{\frac{(1.02)^2}{2}-1.02\right\}-\left\{\frac{(1.01)^2}{2}-1.01\right\}\right]$$

$$= 200\left[\frac{1.0404}{2}-1.02-\frac{1.0201}{2}+1.01\right]$$

$$= 200\left[0.5202-1.02-0.51005+1.01\right]$$

$$= 200\left[0.00015\right] = 0.03.$$

Now, you can try the following exercise.

E4) The life (in hours) X of a certain type of light bulb may be supposed to be a continuous random variable with p.d.f.:

$$f(x) = \begin{cases} \dfrac{A}{x^3}, & 1500 < x < 2500 \\ 0, & \text{elsewhere} \end{cases}$$

Determine the constant A and compute the probability that $1600 \le X \le 2000$.

5.5 DISTRIBUTION FUNCTION

A function F defined for all values of a random variable X by $F(x) = P[X \le x]$ is called the distribution function. It is also known as the cumulative distribution function (c.d.f.) of X since it is the cumulative probability of X up to and including the value x. As X can take any real value, therefore the domain of the distribution function is set of real numbers and as F(x) is a probability value, therefore the range of the distribution function is [0, 1].

Remark 3: Here, X denotes the random variable and x represents a particular value of random variable. $F(x)$ may also be written as $F_X(x)$, which means that it is a distribution function of random variable X.

Discrete Distribution Function

Distribution function of a discrete random variable is said to be discrete distribution function or cumulative distribution function (c.d.f.). Let X be a discrete random variable taking the values $x_1, x_2, x_3, \ldots$ with respective probabilities $p_1, p_2, p_3, \ldots$

Then $F(x_i) = P[X \le x_i] = P[X = x_1] + P[X = x_2] + \ldots + P[X = x_i]$

$$= p_1 + p_2 + \ldots + p_i.$$

The distribution function of X, in this case, is given as in the following table:

X	F(x)
x_1	p_1
x_2	$p_1 + p_2$
x_3	$p_1 + p_2 + p_3$
x_4	$p_1 + p_2 + p_3 + p_4$
.	.
.	.
.	.

The value of F(x) corresponding to the last value of the random variable X is always 1, as it is the sum of all the probabilities. F(x) remains 1 beyond this last value of X also, as it being a probability can never exceed one.

For example, Let X be a random variable having the following probability distribution:

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

Notice that $p(x)$ will be zero for other values of X. Then, Distribution function of X is given by

X	$F(x) = P[X \le x]$
0	$\frac{1}{4}$
1	$\frac{1}{4} + \frac{1}{2} = \frac{3}{4}$
2	$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

Here, for the last value, i.e. for X = 2, we have F(x) =1.

Also, if we take a value beyond 2 say 4, then we get

$$F(4) = P[X \le 4]$$
$$= P[X = 4] + P[X = 3] + P[X \le 2]$$
$$= 0 + 0 + 1 = 1.$$

Example 7: A random variable X has the following probability function:

X	0	1	2	3	4	5	6	7
p(x)	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{100}$	$\frac{1}{50}$	$\frac{17}{100}$

Determine the distribution function of X.

Solution: Here,

F(0) = P[X ≤ 0] = P[X = 0] = 0,

$F(1) = P[X \le 1] = P[X = 0] + P\,[X = 1] = 0 + \frac{1}{10} = \frac{1}{10}$,

$F(2) = P[X \le 2] = P[X = 0] + P\,[X = 1] + [X = 2] = 0 + \frac{1}{10} + \frac{1}{5} = \frac{3}{10}$,

and so on. Thus, the distribution function F(x) of X is given in the following table:

X	F(x) = P[X ≤ x]
0	0
1	$\frac{1}{10}$
2	$\frac{3}{10}$
3	$\frac{3}{10} + \frac{1}{5} = \frac{1}{2}$
4	$\frac{1}{2} + \frac{3}{10} = \frac{4}{5}$
5	$\frac{4}{5} + \frac{1}{100} = \frac{81}{100}$
6	$\frac{81}{100} + \frac{1}{50} = \frac{83}{100}$
7	$\frac{83}{100} + \frac{17}{100} = 1$

Continuous Distribution Function

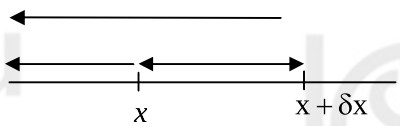
Distribution function of a continuous random variable is called the continuous distribution function or cumulative distribution function (c.d.f.).

Let X be a continuous random variable having the probability density function f(x), as defined in the last section of this unit, then the distribution function F(x) is given by

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx.$$

Also, in the last section, we have defined the p.d.f. f(x) as

$$f(x) = \lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x},$$



$$\therefore f(x) = \lim_{\delta x \to 0} \frac{P[X \le x + \delta x] - P[X \le x]}{\delta x}$$

$$\Rightarrow \; f(x) = \lim_{\delta x \to 0} \frac{F(x + \delta x) - F(x)}{\delta x},$$

$$\Rightarrow f(x) = \text{Derivative of } F(x) \text{ with respect to } x \quad \begin{bmatrix} \text{By definition of} \\ \text{the derivative} \end{bmatrix}$$

$$\Rightarrow f(x) = F'(x)$$

$$\Rightarrow f(x) = \frac{d}{dx}\left(F(x)\right)$$

$$\Rightarrow dF(x) = f(x)\,dx$$

Here, dF(x) is known as the probability differential.

So, $F(x) = \int_{-\infty}^{x} f_X(x)\,dx$ and $F'(x) = f(x)$.

Example 8: The diameter 'X' of a cable is assumed to be a continuous random variable with p.d.f. $f(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$.

Obtain the c.d.f. of X.

Solution: For $0 \le x \le 1$, the c.d.f. of X is given by

$$F(x) = P[X \le x] = \int_0^x f(x)\,dx = \int_0^x 6x(1-x)\,dx$$

$$= 6\int_0^x \left(x - x^2\right)dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^x = 3x^2 - 2x^3$$

$\therefore$ The c.d.f. of X is given by

$$F(x)=\begin{cases} 0, & x<0 \\ 3x^2-2x^3, & 0\le x\le 1. \\ 1, & x>1 \end{cases}$$

Remark 4: In the above example, F(x) is taken as 0 for $x < 0$ since p(x) = 0 for $x < 0$; and F(x) is taken as 1 for $x > 1$ since F(1) = 1 and therefore,

for $x > 1$ also F(x) will remain 1.

Now, you can try the following exercises.

E 5) A random variable X has the following probability distribution:

X	0	1	2	3	4	5	6	7	8
p(x)	k	3k	5k	7k	9k	11k	13k	15k	17k

i) Determine the value of k.

ii) Find the distribution function of X.

E 6) Let X be continuous random variable with p.d.f. given by.

$$f(x)=\begin{cases} \dfrac{x}{2}, & 0\le x<1 \\ \dfrac{1}{2}, & 1\le x<2 \\ \dfrac{1}{2}(3-x), & 2\le x<3 \\ 0, & \text{elsewhere.} \end{cases}.$$

Determine F(x), the c.d.f. of X.

5.6 SUMMARY

Following main points have been covered in this unit of the course:

1) A **random variable** is a function whose domain is a set of possible outcomes and range is a sub-set of the set of reals and has the following properties:

   i) Each particular value of the random variable can be assigned some probability.

   ii) Sum of all the probabilities associated with all the different values of the random variable is unity.

2) A random variable is said to be **discrete random variable** if it has either a finite number of values or a countable number of values, i.e. the values can be arranged in a sequence.

3) If a random variable is such that its values cannot be arranged in a sequence, it is called **continuous random variable**. So, a random variable is said to be continuous if it can take all the possible real (i.e. integer as well as fractional) values between two certain limits.

4) Let X be a discrete r.v. which take on the values x_1, x_2, ... and let $P[X = x_i] = p(x_i)$. The function $p(x_i)$ is called **probability mass function** of X satisfying $p(x_i) \geq 0$ and $\sum_i p(x_i) = 1$. The set $\{(x_1, p(x_1)), (x_2, p(x_2)), ...\}$ specifies the **probability distribution** of discrete r.v. X.

5) Let X be a continuous random variable and $f(x)$ be a continuous function of x. Suppose $(x, x + \delta x)$ be an interval of length δx. Then $f(x)$ defined by $\lim_{\delta x \to 0} \frac{P[x \leq X \leq x + \delta x]}{\delta x} = f(x)$ is called the **probability density function** of X.

   Probability density function has the same properties as that of probability mass function i.e. $f(x) \geq 0$ and $\int_R f(x)\,dx = 1$, where integral has been taken over the entire range R of values of X.

6) A function F defined for all values of a random variable X by $F(x) = P[X \leq x]$ is called the **distribution function**. It is also known as the **cumulative distribution function (c.d.f.)** of X. The domain of the distribution function is a set of real numbers and its range is [0, 1]. Distribution function of a discrete random variable X is said to be **discrete distribution function** and is given by $\{(x_1, F(x_1)), (x_2, F(x_2)), ...\}$. Distribution function of a continuous random variable X having the probability density function $f(x)$ is said to be **continuous distribution function** and is given by

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

   Derivative of $F(x)$ with respect to x is $f(x)$, i.e. $F'(x) = f(x)$.

5.7 SOLUTIONS/ANSWERS

E 1) Let X be the number of bad articles drawn.

$\therefore$ X can take the values 0, 1, 2 with

$P[X = 0] = P[\text{No bad article}]$

= P[Drawing 2 articles from 5 good articles and zero article from 2 bad articles]

$$= \frac{{}^5C_2 \times {}^2C_0}{{}^7C_2} = \frac{5\times4\times1}{7\times6} = \frac{10}{21},$$

P[X = 1] = P[One bad article and 1 good article]

$$= \frac{{}^2C_1 \times {}^5C_1}{{}^7C_2} = \frac{2\times5\times2}{7\times6} = \frac{10}{21}, \text{ and}$$

P[X = 2] = P[Two bad articles and no good article]

$$= \frac{{}^2C_2 \times {}^5C_0}{{}^7C_2} = \frac{1\times1\times2}{7\times6} = \frac{1}{21}$$

$\therefore$ Probability distribution of number of bad articles is:

X	0	1	2
$p(x)$	$\frac{10}{21}$	$\frac{10}{21}$	$\frac{1}{21}$

E2) As $Y = X^2 + 2X$,

$\therefore$ For X = 0, Y = 0 + 0 = 0;

For X = 1, $Y = 1^2 + 2(1) = 3$;

For X = 2, $Y = 2^2 + 2(2) = 8$; and

For X = 3, $Y = 3^2 + 2(3) = 15$.

Thus, the values of Y are 0, 3, 8, 15 corresponding to the values 0, 1, 2, 3 of X and hence

$P[Y = 0] = P[X = 0] = \frac{1}{10}$, $P[Y = 3] = P[X = 1] = \frac{3}{10}$,

$P[Y = 8] = P[X = 2] = \frac{1}{2}$ and $P[Y = 15] = P[X = 3] = \frac{1}{10}$.

$\therefore$ The probability distribution of Y is

Y	0	3	8	15
p(y)	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{10}$

E3) Let X be the number of red balls drawn.

$\therefore$ X can take the values 0, 1, 2, 3.

Let W_i be the event that i^{th} draw gives white ball and R_i be the event that i^{th} draw gives red ball.

$\therefore$ P[X = 0] = P[No Red ball] = $P[W_1 \cap W_2 \cap W_3]$

$= P(W_1).\, P(W_2).\, P(W_3)$

[$\because$ balls are drawn with replacement and hence the draws are independent]

$$= \frac{3}{7} \times \frac{3}{7} \times \frac{3}{7} = \frac{27}{343}$$

P[X = 1] = P[One red and two white]

$$= P\left[\left(R_1 \cap W_2 \cap W_3\right) \text{or} \left(W_1 \cap R_2 \cap W_3\right) \text{or} \left(W_1 \cap W_2 \cap R_3\right)\right]$$

$$= P\left[R_1 \cap W_2 \cap W_3\right] + P\left[W_1 \cap R_2 \cap W_3\right] + P\left[W_1 \cap W_2 \cap R_3\right]$$

$$= P[R_1]P\left[W_2\right]P\left[W_3\right] + P\left[W_1\right]P\left[R_2\right]P\left[W_3\right] + P\left[W_1\right]P\left[W_2\right]P\left[R_3\right]$$

$$= \frac{4}{7} \times \frac{3}{7} \times \frac{3}{7} + \frac{3}{7} \times \frac{4}{7} \times \frac{3}{7} + \frac{3}{7} \times \frac{3}{7} \times \frac{4}{7} = 3 \times \frac{4}{7} \times \frac{3}{7} \times \frac{3}{7} = \frac{108}{343},$$

P[X = 2] = P[Two red and one white]

$$= P\left[\left(R_1 \cap R_2 \cap W_3\right) \text{or } \left(R_1 \cap W_2 \cap R_3\right) \text{or} \left(W_1 \cap R_2 \cap R_3\right)\right]$$

$$= P[R_1]P\left[R_2\right]P\left[W_3\right] + P\left[R_1\right]P\left[W_2\right]P\left[R_3\right] + P\left[W_1\right]P\left[R_2\right]P\left[R_3\right]$$

$$= \frac{4}{7} \times \frac{4}{7} \times \frac{3}{7} + \frac{4}{7} \times \frac{3}{7} \times \frac{4}{7} + \frac{3}{7} \times \frac{4}{7} \times \frac{4}{7} = 3 \times \frac{4}{7} \times \frac{4}{7} \times \frac{3}{7} = \frac{144}{343} .$$

P[X = 3] = P [Three red balls]

$$= P[R_1 \cap R_2 \cap R_3] = P(R_1)\, P(R_2)\, P(R_3) = \frac{4}{7} \times \frac{4}{7} \times \frac{4}{7} = \frac{64}{343}.$$

$\therefore$ Probability distribution of the number of red balls is

X	0	1	2	3
$p(x)$	$\frac{27}{343}$	$\frac{108}{343}$	$\frac{144}{343}$	$\frac{64}{343}$

E4) As $f(x)$ is p.d.f.,

$$\therefore \int_{1500}^{2500} \frac{A}{x^3} dx = 1 \Rightarrow A \int_{1500}^{2500} x^{-3} dx = 1 \Rightarrow A \left[\frac{x^{-2}}{-2}\right]_{1500}^{2500} = 1$$

$$\Rightarrow -\frac{A}{2}\left[\frac{1}{(2500)^2} - \frac{1}{(1500)^2}\right] = 1 \Rightarrow -\frac{A}{20000}\left[\frac{1}{625} - \frac{1}{225}\right] = 1$$

$$\Rightarrow -\frac{A}{20000}\left[\frac{9-25}{5625}\right] = 1 \Rightarrow 16A = 5625 \times 20000$$

$$\Rightarrow \quad A = \frac{5625 \times 20000}{16} = 5625 \times 1250 = 7031250.$$

Now, $P[1600 \le X \le 2000] = \int_{1600}^{2000} f(x) dx = A \int_{1600}^{2000} \frac{1}{x^3} dx$

$$= -\frac{A}{2}\left[\frac{1}{x^2}\right]_{1600}^{2000} = -\frac{A}{2}\left[\frac{1}{(2000)^2} - \frac{1}{(1600)^2}\right]$$

$$= -\frac{A}{20000}\left[\frac{1}{400} - \frac{1}{256}\right] = -\frac{A}{20000}\left[\frac{16-25}{6400}\right]$$

$$= \frac{9 \times 7031250}{20000 \times 6400} = \frac{2025}{4096}$$

E5) i) As the given distribution is probability distribution,

$\therefore$ sum of all the probabilities = 1

$\Rightarrow$ k + 3k + 5k + 7k + 9k + 11k + 13k + 15k + 17k = 1

$\Rightarrow 81\,k = 1 \Rightarrow k = \frac{1}{81}$

ii) The distribution function of X is given in the following table:

X	$F(x) = P[X \le x]$
0	$k = \frac{1}{81}$
1	$k + 3k = 4k = \frac{4}{81}$
2	$4k + 5k = 9k = \frac{9}{81}$
3	$9k + 7k = 16k = \frac{16}{81}$
4	$16k + 9k = 25k = \frac{25}{81}$
5	$25k + 11k = 36k = \frac{36}{81}$
6	$36k + 13k = 49k = \frac{49}{81}$
7	$49k + 15k = 64k = \frac{64}{81}$
8	$64k + 17k = 81k = 1$

E6) For $x < 0$,

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{x} 0\,dx = 0 \quad [\because f(x) = 0 \text{ for } x < 0].$$

For $0 \le x < 1$,

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{x} f(x)\,dx \quad [\because 0 \le x < 1]$$

$$= 0 + \int_{0}^{x} \frac{x}{2}\,dx \qquad \left[\because f(x) = \frac{x}{2} \text{ for } 0 \le x < 1\right]$$

$$= \frac{1}{2}\left[\frac{x^2}{2}\right]_0^x = \frac{x^2}{4}.$$

For $1 \le x < 2$,

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx$$

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

$$= 0 + \int_{0}^{1} \frac{x}{2}\,dx + \int_{1}^{x} \frac{1}{2}\,dx = \frac{1}{4}\left[x^2\right]_0^1 + \frac{1}{2}[x]_1^x$$

$$= \frac{1}{4} + \frac{x}{2} - \frac{1}{2} = \frac{1}{4}(2x-1)$$

For $2 \le x < 3$,

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

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{2} f(x)\,dx + \int_{2}^{x} f(x)\,dx$$

$$= 0 + \int_{0}^{1} \frac{x}{2}\,dx + \int_{1}^{2} \frac{1}{2}\,dx + \int_{2}^{x} \frac{1}{2}(3-x)\,dx$$

$$= \left[\frac{x^2}{4}\right]_0^1 + \frac{1}{2}[x]_1^2 + \frac{1}{2}\left[3x - \frac{x^2}{2}\right]_2^x$$

$$= \frac{1}{4} + \frac{1}{2} + \frac{1}{2}\left[\left(3x - \frac{x^2}{2}\right) - (6-2)\right]$$

$$= \frac{1}{2}\left(3x - \frac{x^2}{2}\right) - \frac{5}{4} = -\frac{x^2}{4} + \frac{3x}{2} - \frac{5}{4}$$

For $3 \le x < \infty$,

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

$$= \int_{-\infty}^{0} f(x)dx + \int_{0}^{1} f(x)dx + \int_{1}^{2} f(x)dx + \int_{2}^{3} f(x)dx \int_{3}^{x} f(x)dx$$

$$= 0 + \int_{0}^{1} \frac{x}{2}dx + \int_{1}^{2} \frac{1}{2}dx + \int_{2}^{3} \frac{1}{2}(3-x)dx + \int_{3}^{x} 0\,dx$$

$$= \left[\frac{x^2}{4}\right]_0^1 + \left[\frac{x}{2}\right]_1^2 + \frac{1}{2}\left[3x - \frac{x^2}{2}\right]_2^3 + 0$$

$$= \frac{1}{4} + \left(1 - \frac{1}{2}\right) + \frac{1}{2}\left[\left(9 - \frac{9}{2}\right) - (6-2)\right]$$

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

$$= \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

Hence, the distribution function is given by:

$$F(x) = \begin{cases} 0, & -\infty < x < 0 \\ \dfrac{x^2}{4}, & 0 \le x < 1 \\ \dfrac{2x-1}{4}, & 1 \le x < 2 \\ -\dfrac{x^2}{4} + \dfrac{3x}{2} - \dfrac{5}{4}, & 2 \le x < 3 \\ 1, & 3 \le x < \infty \end{cases}$$

UNIT 6 BIVARIATE DISCRETE RANDOM VARIABLES

Structure

6.1 INTRODUCTION

In Unit 5, you have studied one-dimensional random variables and their probability mass functions, density functions and distribution functions. There may also be situations where we have to study two-dimensional random variables in connection with a random experiment. For example, we may be interested in recording the number of boys and girls born in a hospital on a particular day. Here, 'the number of boys' and 'the number of girls' are random variables taking the values 0, 1, 2, … and both these random variables are discrete also.

In this unit, we concentrate on the two-dimensional discrete random variables defining them in Sec. 6.2. The joint, marginal and conditional probability mass functions of two-dimensional random variable are described in Sec. 6.3. The distribution function and the marginal distribution function are discussed in Sec. 6.4.

Objectives

A study of this unit would enable you to:

- define two-dimensional discrete random variable;
- specify the joint probability mass function of two discrete random variables;
- obtain the marginal and conditional distributions for two-dimensional discrete random variable;
- define two-dimensional distribution function;
- define the marginal distribution functions; and
- solve various practical problems on bivariate discrete random variables.

6.2 BIVARIATE DISCRETE RANDOM VARIABLES

In Unit 5, the concept of single-dimensional random variable has been studied in detail. Proceeding in analogy with the one-dimensional case, concept of two-dimensional discrete random variables is discussed in the present unit.

A situation where two-dimensional discrete random variable needs to be studied has already been given in Sec. 6.1 of this unit. To describe such situations mathematically, the study of two random variables is introduced.

Definition: Let X and Y be two discrete random variables defined on the sample space S of a random experiment then the function (X, Y) defined on the same sample space is called a two-dimensional discrete random variable. In others words, (X, Y) is a two-dimensional random variable if the possible values of (X, Y) are finite or countably infinite. Here, each value of X and Y is represented as a point (x, y) in the xy-plane.

As an illustration, let us consider the following example:

Let three balls b_1, b_2, b_3 be placed randomly in three cells. The possible outcomes of placing the three balls in three cells are shown in Table 6.1.

Table 6.1 : Possible Outcomes of Placing the Three Balls in Three Cells

Arrangement Number	Placement of the Balls in		
	Cell 1	Cell 2	Cell 3
1	b_1	b_2	b_3
2	b_1	b_3	b_2
3	b_2	b_1	b_3
4	b_2	b_3	b_1
5	b_3	b_1	b_2
6	b_3	b_2	b_1
7	b_1,b_2	b_3	-
8	b_1,b_2	-	b_3
9	-	b_1,b_2	b_3
10	b_1,b_3	b_2	-
11	b_1,b_3	-	b_2
12	-	b_1,b_3	b_2
13	b_2,b_3	b_1	-
14	b_2,b_3	-	b_1
15	-	b_2,b_3	b_1
16	b_1	b_2,b_3	-
17	b_1	-	b_2,b_3
18	-	b_1	b_2,b_3

19	b_2	b_3,b_1	-
20	b_2	-	b_3,b_1
21	-	b_2	b_3,b_1
22	b_3	b_1,b_2	-
23	b_3	-	b_1,b_2
24	-	b_3	b_1,b_2
25	b_1,b_2,b_3	-	-
26	-	b_1,b_2,b_3	-
27	-	-	b_1,b_2,b_3

Now, let X denote the number of balls in Cell 1 and Y be the number of cells occupied. Notice that X and Y are discrete random variables where X take on the values 0, 1, 2, 3 ($\because$ number of balls in Cell 1 may be 0 or 1 or 2 or 3) and Y take on the values 1, 2, 3 ($\because$ number of occupied cells may be 1 or 2 or 3). The possible values of two-dimensional random variable (X, Y), therefore, are all ordered pairs of the values x and y of X and Y, respectively, i.e. are (0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3).

Now, to each possible value $\left(x_i, y_j\right)$ of (X, Y), we can associate a number $p\left(x_i, y_j\right)$ representing $P\left(X = x_i, Y = y_j\right)$ as discussed in the following section of this unit.

6.3 JOINT, MARGINAL AND CONDITIONAL PROBABILITY MASS FUNCTIONS

Let us again consider the example discussed in Sec. 6.2. In this example, we have obtained all possible values of (X, Y), where X is the number of balls in Cell 1 and Y be the number of occupied cells. Now, let us associate numbers $p\left(x_i, y_j\right)$ representing $P\left[X = x_i, Y = y_j\right]$ as follows:

p(0, 1) = P[X = 0, Y = 1] = P[no ball in Cell 1 and 1 cell occupied] = P[Arrangement numbers 26, 27] = $\frac{2}{27}$
p(0, 2) = P[X = 0, Y = 2] = P[no ball in Cell 1 and 2 cells occupied] = P[Arrangement numbers 9, 12, 15, 18, 21, 24] = $\frac{6}{27}$
p(0, 3) = P[X = 0, Y = 3] = P[no ball in Cell 1 and 3 cells occupied] = P[Impossible event] = 0
p(1, 1) = P[X = 1, Y = 1] = P[one ball in Cell 1 and 1 cell occupied] = P[Impossible event] = 0

p(1, 2) = P[X = 1, Y = 2] = P[one ball in Cell 1 and 2 cells occupied] = P[Arrangement numbers 16, 17, 19, 20, 22, 23] $= \frac{6}{27}$
p(1, 3) = P[X = 1, Y = 3] = P[one ball in Cell 1 and 3 cells occupied] = P[Arrangement numbers1 to 6] = $\frac{6}{27}$
p(2, 1) = P[X = 2, Y = 1] = P[two balls in Cell 1 and 1 cell occupied] = P[Impossible event] = 0
p(2, 2) = P[X = 2, Y = 2] = P[two balls in Cell 1 and 2 cells occupied] = P[Arrangement numbers 7, 8, 10, 11, 13, 14] $= \frac{6}{27}$
p(2, 3) = P[X = 2, Y = 3] = P[two balls in Cell 1 and 3 cells occupied] = P[Impossible event] = 0
p(3, 1) = P[X = 3, Y = 1] = P[three balls in Cell 1 and 1 cell occupied] = P[Arrangement number 25] = $\frac{1}{27}$
p(3, 2) = P[X = 3, Y = 2] = P[three balls in Cell 1 and 2 cells occupied] = P[Impossible event] = 0
p(3, 3) = P[X = 3, Y = 3] = P[three balls in Cell 1 and 3 cells occupied] = P[Impossible event] = 0

The values of (X, Y) together with the number associated as above constitute what is known as joint probability distribution of (X, Y) which can be written in the tabular form also as shown below:

X \ Y	1	2	3	Total
0	$2/27$	$6/27$	0	$8/27$
1	0	$6/27$	$6/27$	$12/27$
2	0	$6/27$	0	$6/27$
3	$1/27$	0	0	$1/27$
Total	$3/27$	$18/27$	$6/27$	1

We are now in a position to define joint, marginal and conditional probability mass functions.

Joint Probability Mass Function

Let (X, Y) be a two-dimensional discrete random variable. With each possible outcome (x_i, y_j), we associate a number $p(x_i, y_j)$ representing $P[X = x_i, Y = y_j]$ or $P[X = x_i \cap Y = y_j]$ and satisfying following conditions:

(i) $p(x_i, y_j) \geq 0$

(ii) $\sum_i \sum_j p(x_i, y_j) = 1$

The function p defined for all (x_i, y_j) is in analogy with one-dimensional case and called the **joint probability mass function of X and Y.** It is usually represented in the form of the table as shown in the example discussed above.

Marginal Probability Function

Let (X, Y) be a discrete two-dimensional random variable which take up finite or countably infinite values (x_i, y_j). For each such two-dimensional random variable (X, Y), we may be interested in the probability distribution of X or the probability distribution of Y, individually.

Let us again consider the example of the random placement of three balls in there cells wherein X and Y are the discrete random variables representing "the number of balls in Cell 1" and "the number of occupied cells", respectively. Let us consider Table 6.1 showing the joint distribution of (X, Y). From this table, let us take up the row totals and column totals. The row totals in the table represent the probability distribution of X and the column totals represent the probability distribution of Y, individually. That is,

$$P[X = 0] = \frac{2}{27} + \frac{6}{27} + 0 = \frac{8}{27}$$

$$P[X = 1] = 0 + \frac{6}{27} + \frac{6}{27} = \frac{12}{27}$$

$$P[X = 2] = 0 + \frac{6}{27} + 0 = \frac{6}{27}$$

$$P[X = 3] = \frac{1}{27} + 0 + 0 = \frac{1}{27} \text{ and}$$

$$P[Y = 1] = \frac{2}{27} + 0 + 0 + \frac{1}{27} = \frac{3}{27}$$

$$P[Y = 2] = \frac{6}{27} + \frac{6}{27} + \frac{6}{27} + 0 = \frac{18}{27}$$

$$P[Y = 3] = 0 + \frac{6}{27} + 0 + 0 = \frac{6}{27}$$

These distributions of X and Y, individually, are called the **marginal probability distributions** of X and Y, respectively.

So, if (X, Y) is a discrete two-dimensional random variable which take up the values $\left(x_i, y_j\right)$, then the probability distribution of X is determined as follows:

$$p\left(x_i\right) = P\left[X = x_i\right]$$

$$= P\left[\left(X = x_i \cap Y = y_1\right) \text{or } \left(X = x_i \cap Y = y_2\right) \text{or} \ldots\right]$$

$$= P\left[X = x_i \cap Y = y_1\right] + P\left[X = x_i \cap Y = y_2\right] + P\left[X = x_i \cap Y = y_3\right] + \ldots$$

$$= \sum_j P\left[X = x_i \cap Y = y_j\right]$$

$$= \sum_j p\left(x_i, y_j\right) \left[\begin{array}{l} \because\ p\left(x_i, y_j\right), \text{the joint probability mass} \\ \text{function, is } P\left[X = x_i \cap Y = y_j\right] \end{array}\right]$$

which is known as the marginal probability mass function of X. Similarly, the probability distribution of Y is

$$p\left(y_j\right) = P\left[Y = y_j\right]$$

$$= P\left[X = x_1 \cap Y = y_j\right] + P\left[X = x_2 \cap Y = y_j\right] + \ldots$$

$$= \sum_i P\left[X = x_i \cap Y = y_j\right]$$

$$= \sum_i p\left(x_i, y_j\right)$$

and is known as the marginal probability mass function of Y.

Conditional Probability Mass Function

Let (X, Y) be a discrete two-dimensional random variable. Then the conditional probability mass function of X, given $Y = y$ is defined as

$$p\left(x \mid y\right) = P[X = x \mid Y = y]$$

$$= \frac{P\left[X = x \cap Y = y\right]}{P\left[Y = y\right]}, \text{ provided } P[Y = y] \neq 0$$

$$\left[\because P\left[A \mid B\right] = \frac{P\left[A \cap B\right]}{P\left[B\right]}, P\left(B\right) \neq 0\right]$$

Similarly, the conditional probability mass function of Y, given X = x, is defined as

$$p\left(y \mid x\right) = P\left[Y = y \mid X = x\right] = \frac{P\left[Y = y \cap X = x\right]}{P\left[X = x\right]}$$

Let us again consider the example as already discussed in this section. Suppose, we are interested in finding the conditional probability mass function of X given Y = 2. Then the conditional probabilities are found separately for each value of X given Y = 2. That is, we proceed as follows:

$$P\left[X = 0 \mid Y = 2\right] = \frac{P\left[X = 0 \cap Y = 2\right]}{P\left[Y = 2\right]} = \frac{\frac{6}{27}}{\frac{18}{27}} = \frac{1}{3}$$

$$P\left[X = 1 \mid Y = 2\right] = \frac{P\left[X = 1 \cap Y = 2\right]}{P\left[Y = 2\right]} = \frac{\frac{6}{27}}{\frac{18}{27}} = \frac{1}{3}$$

$$P\left[X = 2 \mid Y = 2\right] = \frac{P\left[X = 2 \cap Y = 2\right]}{P\left[Y = 2\right]} = \frac{\frac{6}{27}}{\frac{18}{27}} = \frac{1}{3}$$

$$P\left[X = 3 \mid Y = 2\right] = \frac{P\left[X = 3 \cap Y = 2\right]}{P\left[Y = 2\right]} = \frac{0}{\frac{18}{27}} = 0$$

[Note that values of numerator and denominator in the above expressions have already been obtained while discussing the joint and marginal probability mass functions in this section of the unit.]

Independence of Random Variables

Two discrete random variables X and Y are said to be independent if and only if

$$P\left[X = x_i \cap Y = y_j\right] = P\left[X = x_i\right] P\left[Y = y_j\right]$$

[$\because$ two events A and B are independent if and only if P(A $\cap$ B) = P(A) P(B)]

Example 1: The following table represents the joint probability distribution of the discrete random variable (X, Y):

X \ Y	1	2
1	0.1	0.2
2	0.1	0.3
3	0.2	0.1

Find **:**

i) The marginal distributions.

ii) The conditional distribution of X given Y = 1.

iii) P[(X + Y) < 4].

Solution:

i) To find the marginal distributions, we have to find the marginal totals, i.e. row totals and column totals as shown in the following table:

Y \ X	1	2	$p(x)$ (Totals)
1	0.1	0.2	0.3
2	0.1	0.3	0.4
3	0.2	0.1	0.3
$p(y)$ (Totals)	0.4	0.6	1

Thus, the marginal probability distribution of X is

X	1	2	3
p(x)	0.3	0.4	0.3

and the marginal probability distribution of Y is

Y	1	2
P(y)	0.4	0.6

ii) As $P[X=1 \mid Y=1]=\dfrac{P[X=1, Y=1]}{P[Y=1]}=\dfrac{0.1}{0.4}=\dfrac{1}{4}$,

$$P[X=2 \mid Y=1]=\frac{P[X=2, Y=1]}{P[Y=1]}=\frac{0.1}{0.4}=\frac{1}{4} \text{ and}$$

$$P[X=3 \mid Y=1]=\frac{P[X=3 \cap Y=1]}{P[Y=1]}=\frac{0.2}{0.4}=\frac{1}{2},$$

$\therefore$ The conditional distribution of X given Y = 1 is

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

(iii) As the values of (X, Y) which satisfy X + Y < 4 are (1, 1), (1, 2) and (2, 1) only.

$$\therefore P\left[(X+Y)<4\right]=P[X=1, Y=1]+P[X=1, Y=2]+P[X=2, Y=1]$$

$$=0.1+0.2+0.1=0.4$$

Example 2: Two discrete random variables X and Y have

$P[X = 0, Y = 0] = \dfrac{2}{9}$, $P[X = 0, Y = 1] = \dfrac{1}{9}$, $P[X=1, Y=0]=\dfrac{1}{9}$, and

$P[X=1, Y=1]=\frac{5}{9}$. Examine whether X and Y are independent?

Solution: Writing the given distribution in tabular form as follows:

X \ Y	0	1	p(x)
0	2/9	1/9	3/9
1	1/9	5/9	6/9
p(y)	3/9	6/9	1

$\therefore\ P[X=0]=\frac{3}{9},\ P[X=1]=\frac{6}{9},$

$P[Y=0]=\frac{3}{9},\ P[Y=1]=\frac{6}{9}$

Now $P[X=0]P[Y=0]=\frac{3}{9}\times\frac{3}{9}=\frac{1}{9}$

But $P[X=0, Y=0]=\frac{2}{9}$

$\therefore\ P[X=0, Y=0]\neq P[X=0]P[Y=0]$

Hence X and Y are not independent

[**Note**: If $P[X=x, Y=y]=P[X=x]P[Y=y]$ for each possible value of X and Y, only then X and Y are independent.]

Here are two exercises for you.

E1) The joint probability distribution of a pair of random variables is given by the following table:

X \ Y	1	2	3
1	1/12	0	1/18
2	1/6	1/9	1/4
3	0	1/5	2/15

i) Evaluate marginal distribution of X.

ii) Evaluate conditional distribution of Y given X = 2

iii) Obtain $P[X+Y<5]$.

E2) For the following joint probability distribution of (X, Y),

X \ Y	1	2	3
1	1/20	1/10	1/10
2	1/20	1/10	1/10
3	1/10	1/10	1/20
4	1/10	1/10	1/20

i) find the probability that Y = 2 given that X = 4,

ii) find the probability that Y = 2, and

iii) examine if the two events X = 4 and Y = 2 are independent.

6.4 JOINT AND MARGINAL DISTRIBUTION FUNCTIONS FOR DISCRETE RANDOM VARIABLES

Two-Dimensional Joint Distribution Function

In analogy with the distribution function $F(x) = P[X \le x]$ of one-dimensional random variable X discussed in Unit 5 of this course, the distribution function of the two-dimensional random variable (X, Y) for all real x and y is defined as

$$F(x, y) = P[X \le x, Y \le y]$$

Marginal Distribution Functions

Let (X, Y) be a two-dimensional discrete random variable having $F(x, y)$ as its distribution function. Now the **marginal distribution function of X** is defined as

$$\begin{aligned} F(x) &= P[X \le x] \\ &= P[X \le x, Y = y_1] + P[X \le x, Y = y_2] + \ldots \\ &= \sum_j P[X \le x, Y = y_j] \end{aligned}$$

Similarly, the **marginal distribution function of Y** is defined as

$$\begin{aligned} F(y) &= P[Y \le y] \\ &= P[X = x_1, Y \le y] + P[X = x_2, Y \le y] + \ldots \\ &= \sum_i P[X = x_i, Y \le y] \end{aligned}$$

Example 3: Considering the probability distribution function given in Example 1, find

i) F(2, 2), F(3,2)

ii) $F_X(3)$

iii) $F_Y(1)$

Solution:

i) $F(2,2) = P[X \le 2, Y \le 2]$

$= P[X = 2, Y \le 2] + P[X = 1, Y \le 2]$

$= P[X = 2, Y = 2] + P[X = 2, Y = 1] + P[X = 1, Y = 2]$

$+ P[X = 1, Y = 1]$

= 0.3 + 0.1 + 0.2 + 0.1 = 0.7

$F(3,2) = P[X \le 3, Y \le 2]$

$= P[X \le 2, Y \le 2] + P[X = 3, Y \le 2]$

$= 0.7 + P[X = 3, Y \le 2]$ $\left[\because \text{first term on R.H.S. has been obtained in part (i) of this example}\right]$

$= 0.7 + P[X = 3, Y = 2] + P[X = 3, Y = 1] = 0.7 + 0.1 + 0.2 = 1$

ii) $F_X(3) = P[X \le 3]$

$= P[X \le 3, Y = 1] + P[X \le 3, Y = 2]$

$= P[X = 3, Y = 1] + P[X = 2, Y = 1] + P[X = 1, Y = 1]$

$+ P[X = 3, Y = 2] + P[X = 2, Y = 2] + P[X = 1, Y = 2]$

= 0.2 + 0.1 + 0.1 +0.1+ 0.3 + 0.2 = 1

iii) $F_Y(1) = P[Y \le 1]$

$= P[X = 1, Y \le 1] + P[X = 2, Y \le 1] + P[X = 3, Y \le 1]$

$= P[X = 1, Y = 1] + P[X = 2, Y = 1] + P[X = 3, Y = 1]$

= 0.1 + 0.1 + 0.2 = 0.4

Example 4: Find the joint and marginal distribution functions for the joint probability distribution given in Example 2.

Solution: For the joint distribution function, we have to find $F(x, y) = P[X \le x, Y \le y]$ for each x and y, i.e. we are to find F(0,0), F(0,1), F(1,0), F(1,1).

$$F(0,0) = P[X \le 0, Y \le 0] = P[X = 0, Y = 0] = \frac{2}{9}$$

$$F(0,1) = P[X \le 0, Y \le 1] = P[X = 0, Y = 0] + P[X = 0, Y = 1]$$

$$= \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$$

$$F(1,0) = P[X \le 1, Y \le 0] = P[X = 1, Y = 0] + P[X = 0, Y = 0]$$

$$= \frac{1}{9} + \frac{2}{9} = \frac{3}{9}$$

$$F(1,1) = P[X \le 1, Y \le 1] = P[X = 1, Y = 1] + P[X = 1, Y = 0]$$

$$+ P[X = 0, Y = 1] + P[X = 0, Y = 0]$$

$$= \frac{5}{9} + \frac{1}{9} + \frac{1}{9} + \frac{2}{9} = 1$$

Above distribution function $F(x, y)$ can be shown in the tabular form as follows:

	$Y \le 0$	$Y \le 1$
$X \le 0$	$2/9$	$3/9$
$X \le 1$	$3/9$	1

Marginal distribution function of X is obtained on finding $F(x) = P[X \le x]$ for each x, i.e. we have to obtain $F_X(0)$, $F_X(1)$.

$$F_X(0) = P[X \le 0] = P[X = 0]$$

$$= P[X = 0, Y = 0] + P[X = 0, Y = 1]$$

$$= \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$$

$$F_X(1) = P[X \le 1] = P[X \le 1, Y = 0] + P[X \le 1, Y = 1]$$

$$= P[X = 1, Y = 0] + P[X = 0, Y = 0]$$

$$+ P[X = 1, Y = 1] + P[X = 0, Y = 1]$$

$$= \frac{1}{9} + \frac{2}{9} + \frac{5}{9} + \frac{1}{9} = 1$$

$\therefore$ marginal distribution function of X is given as

X	F(x)
≤ 0	$3/9$
≤ 1	1

Similarly, marginal distribution function of Y can be obtained. [Do it yourself]

Here is an exercise for you.

E3) Obtain the joint and marginal distribution functions for the joint probability distribution given in **E 1).**

Now before ending this unit, let us summarizes what we have covered in it.

6.5 SUMMARY

In this unit we have covered the following main points:

1) If X and Y be two discrete random variables defined on the sample space S of a random experiment then the function (X, Y) defined on the same sample space is called a **two-dimensional discrete random variable**. In others words, (X, Y) is a two-dimensional random variable if the possible values of (X, Y) are finite or countably infinite.

2) A number $p(x_i, y_j)$ associated with each possible outcome (x_i, y_j) of a two-dimensional discrete random variable (X, Y) is called the **joint probability mass function of X and Y** if it satisfies the following conditions:

   (i) $p(x_i, y_j) \geq 0$

   (ii) $\sum_i \sum_j p(x_i, y_j) = 1$

3) If (X,Y) is a discrete two-dimensional random variable which take up the values (x_i, y_j), then the probability distribution of X given by $p(x_i) = \sum_j p(x_i, y_j)$ is known as the **marginal probability mass function of X** and the probability distribution of Y given by $p(y_j) = \sum_i p(x_i, y_j)$ is known as the **marginal probability mass function of Y**.

4) The **conditional probability mass function of X given** $Y = y$ in case of a two-dimensional discrete random variable (X, Y) is defined as

$$p(x \mid y) = P[X = x \mid Y = y]$$

$$= \frac{P[X = x \cap Y = y]}{P[Y = y]}; \text{ and}$$

   the **conditional probability mass function of Y, given X = x** is defined as

$$p(y \mid x) = P[Y = y \mid X = x]$$

$$= \frac{P[Y = y \cap X = x]}{P[X = x]}$$

5) Two discrete random variables X and Y are said to be independent if and only if

$$P[X = x_i \cap Y = y_j] = P[X = x_i] P[Y = y_j]$$

6.6 SOLUTIONS/ANSWERS

E1) Let us compute the marginal totals. Thus, the complete table with marginal totals is given as

X \ Y	1	2	3	$p(x)$
1	$\frac{1}{12}$	0	$\frac{1}{18}$	$\frac{1}{12}+0+\frac{1}{18}=\frac{5}{36}$
2	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{4}$	$\frac{1}{6}+\frac{1}{9}+\frac{1}{4}=\frac{19}{36}$
3	0	$\frac{1}{5}$	$\frac{2}{15}$	$0+\frac{1}{5}+\frac{2}{15}=\frac{1}{3}$
$p(y)$	$\frac{1}{4}$	$\frac{14}{45}$	$\frac{79}{180}$	1

Therefore,

i) Marginal distribution of X is

X	$p(x)$
1	$5/36$
2	$19/36$
3	$1/3$

ii) $$P[Y=1 \mid X=2] = \frac{P[Y=1, X=2]}{P[X=2]} = \frac{1}{6}\times\frac{36}{19} = \frac{6}{19}$$

$$P[Y=2 \mid X=2] = \frac{P[Y=2, X=2]}{P[X=2]} = \frac{1}{9}\times\frac{36}{19} = \frac{4}{19}$$

$$P[Y=3 \mid X=2] = \frac{P[Y=3, X=2]}{P[X=2]} = \frac{1}{4}\times\frac{36}{19} = \frac{9}{19}$$

$\therefore$ The conditional distribution of Y given X = 2 is

Y	$P[Y=y \mid X=2]$
1	$6/19$
2	$4/19$
3	$9/19$

iii) $P[X+Y<5]$

$=P[X=1, Y=1]+P[X=1, Y=2]+P[X=1, Y=3]$

$+P[X=2, Y=1]+P[X=2, Y=2]+P[X=3, Y=1]$

$=\frac{1}{12}+0+\frac{1}{18}+\frac{1}{6}+\frac{1}{9}+0=\frac{15}{36}$.

E2) First compute the marginal totals, then you will be able to find

i) $P[X=4]=\frac{1}{4}$, and hence

$$P[Y=2 \mid X=4]=\frac{P[Y=2,\ X=4]}{P[X=4]}=\frac{2}{5}$$

ii) $P[Y=2]=\frac{2}{5}$

iii) $P[X=4, Y=2]=\frac{1}{10},\ \ P[X=4]=\frac{1}{4}, P[Y=2]=\frac{2}{5}$

$P[X = 4]\ P[Y = 2] = \frac{1}{4}\times\frac{2}{5} = \frac{1}{10}$

$\therefore$ X= 4 and Y= 2 are independent

E3) To obtain joint distribution function $F(x, y)=P[X \leq x, Y \leq y]$, we have to obtain

$F(x, y)$ for each value of X and Y, i.e. we have to obtain

$F(1,1), F(1,2), F(1,3), F(2,1), F(2,2), F(2,3), F(3,1), F(3,2), F(3,3)$.

Then, the distribution function in tabular form is

	$Y \leq 1$	$Y \leq 2$	$Y \leq 3$
$X \leq 1$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{5}{36}$
$X \leq 2$	$\frac{1}{4}$	$\frac{13}{36}$	$\frac{2}{3}$
$X \leq 3$	$\frac{1}{4}$	$\frac{101}{180}$	1

Marginal distribution function of X is given as

X	F(x)
≤ 1	5/36
≤ 2	2/3
≤ 3	1

Marginal distribution function of Y is

Y	F(y)
≤ 1	1/4
≤ 2	101/180
≤ 3	1

UNIT 7 BIVARIATE CONTINUOUS RANDOM VARIABLES

Structure

7.1 INTRODUCTION

In Unit 6, we have defined the bivariate discrete random variable (X, Y), where X and Y both are discrete random variables. It may also happen that one of the random variables is discrete and the other is continuous. However, in most applications we deal only with the cases where either both random variables are discrete or both are continuous. The cases where both random variables are discrete have already been discussed in Unit 6. Here, in this unit, we are going to discuss the cases where both random variables are continuous.

In Unit 6, you have studied the joint, marginal and conditional probability functions and distribution functions in context of bivariate discrete random variables. Similar functions, but in context of bivariate continuous random variables, are discussed in this unit.

Bivariate continuous random variable is defined in Sec. 7.2. Joint and marginal density functions are described in Sec. 7.3. Sec. 7.4 deals with the conditional distribution and density functions. Independence of two continuous random variables is dealt with in Sec. 7.5. Some practical problems on two-dimensional continuous random variables are taken up in Sec. 7.6.

Objectives

A study of this unit would enable you to:

- define two-dimensional continuous random variable;
- specify the joint and marginal probability density functions of two continuous random variables;
- obtain the conditional density and distribution functions for two-dimensional continuous random variable;
- check the independence of two continuous random variables; and
- solve various practical problems on bivariate continuous random variables.

7.2 BIVARIATE CONTINUOUS RANDOM VARIABLES

Definition: If X and Y are continuous random variables defined on the sample space S of a random experiment, then (X, Y) defined on the same sample space S is called bivariate continuous random variable if (X, Y) assigns a point in xy-plane defined on the sample space S. Notice that it (unlike discrete random variable) assumes values in some non-countable set. Some examples of bivariate continuous random variable are:

1. A gun is aimed at a certain point (say origin of the coordinate system). Because of the random factors, suppose the actual hit point is any point (X, Y) in a circle of radius unity about the origin.

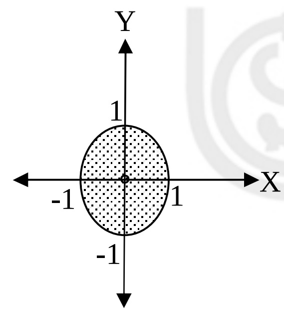


Fig.7.1: Actual Hit Point when a Gun is Aimed at a Certain Point

Then (X, Y) assumes all the values in the circle $\{(x, y): x^2 + y^2 \le 1\}$ i.e. (X, Y) assumes all values corresponding to each and every point in the circular region as shown in Fig.7.1. Here, (X, Y) is bivariate continuous random variable.

2. (X, Y) assuming all values in the rectangle $\{(x, y): a \le x \le b, c \le y \le d\}$ is a bivariate continuous random variable.

   Here, (X, Y) assumes all values corresponding to each and every point in the rectangular region as shown in Fig.7.2.

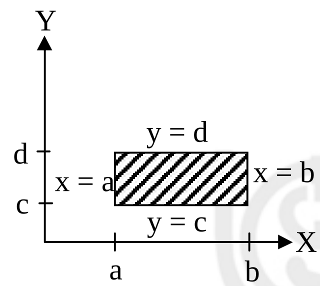


Fig.7.2: (X, Y) Assuming All Values in the Rectangle $\{(x,y): a \le x \le b, c \le y \le d\}$

3. In a statistical survey, let X denotes the daily number of hours a child watches television and Y denotes the number of hours he/she spends on the studies. Here, (X, Y) is a two-dimensional continuous random variable.

7.3 JOINT AND MARGINAL DISTRIBUTION AND DENSITY FUNCTIONS

Two-Dimensional Continuous Distribution Function

The distribution function of a two-dimensional continuous random variable (X, Y) is a real-valued function and is defined as

$F(x, y) = P[X \le x, Y \le y]$ for all real x and y.

Notice that the above function is in analogy with one-dimensional continuous random variable case as studied in Unit 5 of the course.

Remark 1: $F(x, y)$ can also be written as $F_{X,Y}(x, y)$.

Joint Probability Density Function

Let (X, Y) be a continuous random variable assuming all values in some region R of the xy-plane. Then, a function $f(x, y)$ such that

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\,dy\,dx$$

is defined to be a joint probability density function.

As in the one-dimensional case, a joint probability density function has the following properties.

i) $f(x, y) \ge 0$

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

Remark 2:

As in the one-dimensional case, $f(x, y)$ does not represent the probability of anything. However, for positive δx and δy sufficiently small, $f(x, y)\delta x \delta y$ is approximately equal to

$P[x \le X \le x + \delta x, y \le Y \le y + \delta y]$.

In the one-dimensional case, you have studied that for positive δx sufficiently small $f(x)\delta x$ is approximately equal to $P[x \le X \le x + \delta x]$. So, the two-dimensional case is in analogy with the one-dimensional case.

Remark 3:

In analogy with the one-dimensional case [See Sec. 5.4 of Unit 5 of this course],

$f(x, y)$ can be written as $\lim\limits_{\substack{\delta x \to 0 \\ \delta y \to 0}} \dfrac{P[x \le X \le x + \delta x, y \le Y \le y + \delta y]}{\delta x \delta y}$

and is equal to

$\dfrac{\partial^2}{\partial x \partial y}(F(x, y))$, i.e. second order partial derivative with respect to x and y.

[See Sec. 5.5 of Unit 5 where $f(x) = \frac{d}{dx}(F(x))$]

Note: $\frac{\partial^2}{\partial x \partial y}(F(x, y))$ means first differentiate $F(x, y)$ partially w.r.t. y and then the resulting function w.r.t. x. When we differentiate a function partially w.r.t. one variable, then the other variable is treated as constant

For example, Let $F(x, y) = xy^3 + x^2 y$

If we differentiate it partially w.r.t. y, we have

$\frac{\partial}{\partial y}(F(x, y)) = x(3y^2) + x^2.1$ [$\because$ here, x is treated as constant.]

If we now differentiate this resulting expression w.r.t. x, we have

$\frac{\partial^2}{\partial x \partial y}(F(x, y)) = 3y^2 + 2x$ [$\because$ here, y is treated as constant.]

Marginal Continuous Distribution Function

Let (X, Y) be a two-dimensional continuous random variable having $f(x, y)$ as its joint probability density function. Now, the marginal distribution function of the continuous random variable X is defined as

$$F(x) = P[X \le x]$$

$$= P[X \le x, y < \infty] \quad [\because \text{ for } X \le x, Y \text{ can take any real value}]$$

$$= \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(x, y) dy \right] dx,$$

and the marginal distribution function of the continuous random variable Y is defined as

$$F(y) = P[Y \le y]$$

$$= P[Y \le y, X < \infty] \quad [\because \text{ for } Y \le y, X \text{ can take any real value}]$$

$$= \int_{-\infty}^{y} \left[\int_{-\infty}^{\infty} f(x, y) dx \right] dy$$

Marginal Probability Density Functions

Let (X, Y) be a two-dimensional continuous random variable having $F(x, y)$ and $f(x, y)$ as its distribution function and joint probability density function, respectively. Let $F(x)$ and $F(y)$ be the marginal distribution functions of X and Y, respectively. Then, the marginal probability density function of X is given as

$$f(x) = \int_{-\infty}^{\infty} f(x, y)\,dy,$$

or, it may also be obtained as $\frac{d}{dx}\left(F(x)\right)$,

and the marginal probability density function of Y is given as

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

or

$$= \frac{d}{dy}\left(F(y)\right)$$

7.4 CONDITIONAL DISTRIBUTION AND DENSITY FUNCTIONS

Conditional Probability Density Function

Let (X, Y) be a two-dimensional continuous random variable having the joint probability density function f(x, y). The conditional probability density function of Y given X = x is defined as

$f(y \mid x) = \frac{f(x, y)}{f(x)}$, where f(x) > 0 is the marginal density of X.

Similarly, the conditional probability density function of X given Y = y is defined to be

$f(x \mid y) = \frac{f(x, y)}{f(y)}$, where f(y) > 0 is the marginal density of Y.

As $f(y \mid x)$ and $f(x \mid y)$, though conditional yet, are the probability density functions, hence possess the properties of a probability density function. Properties of $f(y \mid x)$ are:

i) $f(y \mid x)$ is clearly ≥ 0

ii) $$\int_{-\infty}^{\infty} f(y \mid x)\,dy = \int_{-\infty}^{\infty} \frac{f(x, y)}{f(x)}\,dy$$

$$= \frac{1}{f(x)}\left[\int_{-\infty}^{\infty} f(x, y)\,dy\right]$$

$$= \frac{1}{f(x)}\left[f(x)\right] \qquad \left[\because \int_{-\infty}^{\infty} f(x, y)\,dy \text{ is the marginal probability density function of X}\right]$$

$$= 1$$

Similarly, $f(x \mid y)$ satisfies

i) $f(x \mid y) \geq 0$ and

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

Conditional Continuous Distribution Function

For a two-dimensional continuous random variable (X, Y), the conditional distribution function of Y given X = x is defined as

$$F(y \mid x) = P[Y \leq y \mid X = x]$$

$$= \int_{-\infty}^{y} f(y \mid x)\,dy \text{, for all x such that } f(x) > 0;$$

and the conditional distribution function of X given Y = y is defined as

$$F(x \mid y) = P[X \leq x \mid Y = y]$$

$$= \int_{-\infty}^{x} f(x \mid y)\,dx \text{, for all y such that f(y) > 0.}$$

7.5 STOCHASTIC INDEPENDENCE OF TWO CONTINUOUS RANDOM VARIABLES

You have already studied in Unit 3 of this course that independence of events is closely related to conditional probability, i.e. if events A and B are independent, then $P[A \mid B] = P[A]$, i.e. conditional probability of A is equal to the unconditional probability of A. Likewise independence of random variables is closely related to conditional distributions of random variables, i.e. two random variables X and Y with joint probability function $f(x, y)$ and marginal probability functions $f(x)$ and $f(y)$ respectively are said to be stochastically independent if and only if

i) $f(y \mid x) = f(y)$

ii) $f(x \mid y) = f(x)$.

Now, as defined in Sec. 7.4, we have

$$f(y \mid x) = \frac{f(x, y)}{f(x)}$$

$\Rightarrow f(x, y) = f(x) f(y \mid x)$ [On cross-multiplying]

So, if X and Y are independent, then

$f(x, y) = f(x) f(y)$ [$\because f(y \mid x) = f(y)$]

Remark 4: The random variables, if independent, are actually stochastically independent but often the word "stochastically" is omitted.

Definition: Two random variables are said to be (stochastically) independent if and only if their joint probability density function is the product of their marginal density functions.

Let us now take up some problems on the topics covered so far in this unit.

7.6 PROBLEMS ON TWO-DIMENSIONAL CONTINUOUS RANDOM VARIABLES

Example 1: Let X and Y be two random variables. Then for

$$f(x,y)=\begin{cases} k(2x+y), & 0<x<1,\ 0<y<2 \\ 0, & \text{elsewhere} \end{cases}$$

to be a joint density function, what must be the value of k ?

Solution: As $f(x,y)$ is the joint probability density function,

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

$$\Rightarrow \int_0^1\int_0^2 f(x,y)\,dy\,dx=1 \qquad [\because 0<x<1,\ 0<y<2\ ]$$

$$\Rightarrow \int_0^1\int_0^2 k(2x+y)\,dy\,dx=1$$

$$\Rightarrow k\int_0^1\left[\int_0^2 (2x+y)\,dy\right]dx=1$$

$$\Rightarrow k\int_0^1\left[2xy+\frac{y^2}{2}\right]_0^2 dx=1$$

[Firstly the integral has been done w.r.t. y treating x as constant.]

$$\Rightarrow k\int_0^1\left[2x(2)+\frac{(2)^2}{2}-0\right]dx=1$$

$$\Rightarrow k\int_0^1 (4x+2)\,dx=1$$

$$\Rightarrow k\left[\frac{4x^2}{2}+2x\right]_0^1=1$$

$$\Rightarrow k\left[\frac{4}{2}+2-0\right]=1 \ \Rightarrow 4k=1 \Rightarrow k=\frac{1}{4}$$

Example 2: Let the joint density function of a two-dimensional random variable (X, Y) be:

$$f(x,y)=\begin{cases} x+y & \text{for } 0\le x<1 \text{ and } 0\le y<1 \\ 0, & \text{otherwise}\end{cases}$$

Find the conditional density function of Y given X.

Solution: The conditional density function of Y given X is $f(y\,|\,x)=\dfrac{f(x,y)}{f(x)}$,

where $f(x,y)$ is the joint density function, which is given; and $f(x)$ is the marginal density function which, by definition, is given by

$$f(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$$

$$=\int_0^1 f(x,y)\,dy \qquad [\because\ 0\le y<1]$$

$$=\int_0^1 (x+y)\,dy$$

$$=\left[xy+\frac{y^2}{2}\right]_0^1.$$

$$=\left[x(1)+\frac{(1)^2}{2}-0\right]=x+\frac{1}{2},\ 0\le x<1.$$

$\therefore$ the conditional density function of Y given X is

$$f(y\,|\,x)=\frac{f(x,y)}{f(x)}=\frac{x+y}{x+\frac{1}{2}},\ \text{for } 0\le x<1 \text{ and } 0\le y<1.$$

Example 3: Two-dimensional random variable (X, Y) have the joint density

$$f(x,y)=\begin{cases} 8xy, & 0<x<y<1 \\ 0, & \text{otherwise}\end{cases}$$

i) Find $P[X<\frac{1}{2}\cap Y<\frac{1}{4}]$.

ii) Find the marginal and conditional distributions.

iii) Are X and Y independent?

Solution:

i) $$P\left[X<\frac{1}{2}\cap Y<\frac{1}{4}\right]=\int_0^{\frac{1}{2}}\int_0^{\frac{1}{4}} f(x,y)\,dy\,dx=\int_0^{\frac{1}{2}}\int_0^{\frac{1}{4}} 8xy\,dy\,dx=\int_0^{\frac{1}{2}} 8x\left[\frac{y^2}{2}\right]_0^{\frac{1}{4}} dx$$

$$= \int_0^{\frac{1}{2}} 8x \left[\frac{1}{16(2)}\right] dx \; = \int_0^{\frac{1}{2}} \frac{x}{4} dx \; = \frac{1}{4}\left[\frac{x^2}{2}\right]_0^{\frac{1}{2}}$$

$$= \frac{1}{4}\left[\frac{1}{8}\right] = \frac{1}{32}.$$

ii) Marginal density function of X is

$$f(x) = \int_x^1 f(x,y)\,dy \qquad [\because 0 < x < y < 1]$$

$$= \int_x^1 8xy\,dy \; = 8x\left[\frac{y^2}{2}\right]_x^1$$

$$= 8x\left[\frac{1}{2} - \frac{x^2}{2}\right] = 4x\left(1 - x^2\right) \text{ for } 0 < x < 1$$

Marginal density function of Y is

$$f(y) = \int_0^y f(x,y)\,dx \qquad [\because 0 < x < y]$$

$$= \int_0^y 8xy\,dx$$

$$= 8y\left[\frac{x^2}{2}\right]_0^y = \frac{8y^3}{2} = 4y^3 \quad \text{for} \quad 0 < y < 1$$

Conditional density function of X given Y(0 < Y < 1) is

$$f(x \mid y) = \frac{f(x,y)}{f(y)}$$

$$= \frac{8xy}{4y^3} = \frac{2x}{y^2},\; 0 < x < y$$

Conditional density function of Y given X(0 < X < 1) is

$$f(y \mid x) = \frac{f(x,y)}{f(x)}$$

$$= \frac{8xy}{4x\left(1-x^2\right)} \; = \frac{2y}{\left(1-x^2\right)},\; x < y < 1$$

iii) $f(x,y) = 8xy,$

But $f(x)f(y) = 4x\left(1-x^2\right)4y^3$

$$= 16x\left(1-x^2\right)y^3$$

$\therefore f(x, y) \neq f(x) f(y)$

Hence, X and Y are not independent random variables.

Now, you can try some exercises.

E1) Let X and Y be two random variables. Then for

$$f(x, y) = \begin{cases} kxy & \text{for } 0 < x < 4 \text{ and } 1 < y < 5 \\ 0, & \text{otherwise} \end{cases}$$

to be a joint density function, what must be the value of k?

E2) If the joint p.d.f. of a two-dimensional random variable (X, Y) is given by

$$f(x, y) = \begin{cases} 2 & \text{for } 0 < x < 1 \text{ and } 0 < y < x \\ 0, & \text{otherwise,} \end{cases}$$

Then,

i) Find the marginal density functions of X and Y.

ii) Find the conditional density functions.

iii) Check for independence of X and Y.

E3) If (X, Y) be two-dimensional random variable having joint density function.

$$f(x, y) = \begin{cases} \frac{1}{8}(6 - x - y); & 0 < x < 2, 2 < y < 4 \\ 0, & \text{elsewhere} \end{cases}$$

Find (i) $P[X < 1, Y < 3]$ (ii) $P[X < 1 | Y < 3]$

Now before ending this unit, let's summarize what we have covered in it.

7.7 SUMMARY

In this unit, we have covered the following main points:

1) If X and Y are continuous random variables defined on the sample space S of a random experiment, then (X, Y) defined on the same sample space S is called **bivariate continuous random variable** if (X, Y) assigns a point in xy-plane defined on the sample space S.

2) The distribution function of a two-dimensional continuous random variable (X, Y) is a real-valued function and is defined as

$$F(x, y) = P[X \leq x, Y \leq y] \text{ for all real x and y.}$$

3) A function $f(x, y)$ is called **joint probability density function** if it is such that

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dy dx$$

and satisfies

i) $f(x, y) \geq 0$

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

4) The **marginal distribution function** of the continuous random variable X is defined as

$$F(x) = P[X \leq x] = \int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(x, y)\,dy\right]dx,$$

and that of continuous random variable Y is defined as

$$F(y) = P[Y \leq y] = \int_{-\infty}^{y}\left[\int_{-\infty}^{\infty} f(x, y)\,dx\right]dy.$$

5) The **marginal probability density function** of X is given as

$$f(x) = \int_{-\infty}^{\infty} f(x, y)\,dy = \frac{d}{dx}(F(x)),$$

and that of Y is given as

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

6) The **conditional probability density function** of Y given X = x is defined as

$$f(y \mid x) = \frac{f(x, y)}{f(x)},$$

and that of X given Y = y is defined as

$$f(x \mid y) = \frac{f(x, y)}{f(y)}.$$

7) The **conditional distribution function** of Y given X = x is defined as

$$F(y \mid x) = \int_{-\infty}^{y} f(y \mid x)\,dy, \text{ for all x such that } f(x) > 0;$$

and that of X given Y = y is defined as

$$F(x \mid y) = \int_{-\infty}^{x} f(x \mid y)\,dx, \text{ for all y such that f(y) > 0.}$$

8) Two random variables are said to be **(stochastically) independent** if and only if their joint probability density function is the product of their marginal density functions.

7.8 SOLUTIONS/ANSWERS

E1) As $f(x, y)$ is the joint probability density function,

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

$$\Rightarrow \int_{0}^{4}\int_{1}^{5} kxy\,dy\,dx = 1 \Rightarrow k\int_{0}^{4}\left[\int_{1}^{5} xy\,dy\right]dx = 1$$

$$\Rightarrow k\int_{0}^{4}\left[x\frac{y^2}{2}\right]_1^5 dx = 1 \Rightarrow k\int_{0}^{4} 12x\,dx = 1$$

$$\Rightarrow 12k\left[\frac{x^2}{2}\right]_0^4 = 1 \Rightarrow 96\,k = 1$$

$$\Rightarrow k = \frac{1}{96}$$

E2) i) Marginal density function of Y is given by

$$f(y) = \int_{-\infty}^{\infty} f(x, y)\,dx = \int_{y}^{1} 2dx$$

[As x is involved in both the given ranges, i.e. $0 < x < 1$ and $0 < y < x$; therefore, here we will combine both these intervals and hence have $0 < y < x < 1$. $\therefore$ x takes the values from y to 1]

$$= [2x]_y^1 = 2 - 2y$$

$= 2 – 2y$

$= 2(1– y),\ 0 < y < 1$

Marginal density function of X is given by

$$f(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$$

$$= \int_{0}^{x} 2dy \qquad [\because 0 < \underbrace{y < x}\ < 1]$$

$$= 2[y]_0^x$$

$$= 2x,\ 0 < x < 1.$$

ii) Conditional density function of Y given $X(0 < X < 1)$ is

$$f(y \mid x) = \frac{f(x, y)}{f(x)} = \frac{2}{2x} = \frac{1}{x};\ 0 < y < x$$

Conditional density function of X and given $Y(0 < Y < 1)$ is

$$f(x \mid y) = \frac{f(x, y)}{f(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y},\ y < x < 1$$

iii) $f(x, y) = 2,$

$$f(x)f(y) = 2(2x)(1-y)$$

As $f(x, y) \neq f(x)f(y),$

$\therefore$ X and Y are not independent.

E3) (i) $P[X < 1, Y < 3] = \int\limits_{-\infty}^{1}\int\limits_{-\infty}^{3} f(x, y)\,dy\,dx$

$$= \int_0^1\int_2^3 \frac{1}{8}(6 - x - y)\,dy\,dx$$

$$= \int_0^1 \left[\frac{1}{8}\left(6y - xy - \frac{y^2}{2}\right)\right]_2^3 dx$$

$$= \frac{1}{8}\int_0^1 \left[\left\{6(3) - x(3) - \frac{9}{2}\right\} - \{12 - 2x - 2\}\right] dx$$

$$= \frac{1}{8}\int_0^1 \left[\left(18 - 3x - \frac{9}{2}\right) - (10 - 2x)\right] dx$$

$$= \frac{1}{8}\int_0^1 \left(\frac{7}{2} - x\right) dx \;\; = \frac{1}{8}\left[\frac{7}{2}x - \frac{x^2}{2}\right]_0^1 \;\; = \frac{1}{8}\left[\frac{7}{2} - \frac{1}{2}\right] = \frac{3}{8}$$

ii) $P[X < 1 \mid Y < 3] = \dfrac{P[X < 1, Y < 3]}{P[Y < 3]}$

where $P(Y < 3) = \int_0^2\int_2^3 \frac{1}{8}(6 - x - y)\,dy\,dx$

$$= \frac{1}{8}\int_0^2 \left[6y - xy - \frac{y^2}{2}\right]_2^3 dx\,.$$

$$= \frac{1}{8}\int_0^2 \left[\left\{18 - 3x - \frac{9}{2}\right\} - \{12 - 2x - 2\}\right] dx$$

$$= \frac{1}{8}\int_0^2 \left[\left(18 - 3x - \frac{9}{2}\right) - (10 - 2x)\right] dx$$

$$= \frac{1}{8}\int_0^2 \left(\frac{7}{2} - x\right) dx$$

$$= \frac{1}{8}\left[\frac{7}{2}x - \frac{x^2}{2}\right]_0^2$$

$$= \frac{1}{8}\left[7 - \frac{4}{2} - 0\right]$$

$$= \frac{5}{8}$$

$$\therefore P\left[X < 1 \middle| Y < 3\right] = \frac{3/8}{5/8} \quad \left[\begin{array}{l}\because \text{value of numerator is} \\ \text{already calculated in part(i)}\end{array}\right]$$

$$= \frac{3}{5}$$

UNIT 8 MATHEMATICAL EXPECTATION

Structure

8.1 INTRODUCTION

In Units 1 to 4 of this course, you have studied probabilities of different events in various situations. Concept of univariable random variable has been introduced in Unit 5 whereas that of bivariate random variable in Units 6 and 7. Before studying the present unit, we advice you to go through the above units.

You have studied the methods of finding mean, variance and other measures in context of frequency distributions in MST-002 (Descriptive Statistics). Here, in this unit we will discuss mean, variance and other measures in context of probability distributions of random variables. Mean or Average value of a random variable taken over all its possible values is called the expected value or the expectation of the random variable. In the present unit, we discuss the expectations of random variables and their properties.

In Secs. 8.2, 8.3 and 8.4, we deal with expectation and its properties. Addition and multiplication laws of expectation have been discussed in Sec. 8.5.

Objectives

After studying this unit, you would be able to:

- find the expected values of random variables;
- establish the properties of expectation;
- obtain various measures for probability distributions; and
- apply laws of addition and multiplication of expectation at appropriate situations.

8.2 EXPECTATION OF A RANDOM VARIABLE

In Unit 1 of MST-002, you have studied that the mean for a frequency distribution of a variable X is defined as

$$\text{Mean} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}.$$

If the frequency distribution of the variable X is given as

$$\begin{array}{lllll} x: & x_1 & x_2 & x_3 \ldots x_n \\ f: & f_1 & f_2 & f_3 \ldots f_n \end{array}$$

The above formula of finding mean may be written as

$$\text{Mean} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{f_1x_1 + f_2x_2 + \ldots + f_nx_n}{\sum_{i=1}^{n} f_i}$$

$$= \frac{x_1f_1}{\sum_{i=1}^{n} f_i} + \frac{x_2f_2}{\sum_{i=1}^{n} f_i} + \ldots + \frac{x_nf_n}{\sum_{i=1}^{n} f_i}$$

$$= x_1\left(\frac{f_i}{\sum_{i=1}^{n} f_i}\right) + x_2\left(\frac{f_2}{\sum_{i=1}^{n} f_i}\right) + \ldots + x_n\left(\frac{f_n}{\sum_{i=1}^{n} f_i}\right)$$

Notice that $\frac{f_1}{\sum_{i=1}^{n} f_i}, \frac{f_2}{\sum_{i=1}^{n} f_i}, \ldots, \frac{f_n}{\sum_{i=1}^{n} f_i}$ are, in fact, the relative frequencies or the proportion of individuals corresponding to the values $x_1, x_2, \ldots, x_n$ respectively of variable X and hence can be replaced by probabilities. [See Unit 2 of this course]

Let us now define a similar measure for the probability distribution of a random variable X which assumes the values say $x_1, x_2, \ldots, x_n$ with their associated probabilities $p_1, p_2, \ldots, p_n$. This measure is known as expected value of X and in the similar way is given as

$x_1(p_1) + x_2(p_2) + \ldots + x_n(p_n) = \sum_{i=1}^{n} x_i p_i$ with only difference is that the role of relative frequencies has now been taken over by the probabilities. The expected value of X is written as E(X).

The above aspect can be viewed in the following way also:

Mean of a frequency distribution of X is $\dfrac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$, similarly mean of a probability distribution of r.v. X is $\dfrac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i}$.

Now, as we know that $\sum_{i=1}^{n} p_i = 1$ for a probability distribution, therefore the mean of the probability distribution becomes $\sum_{i=1}^{n} x_i p_i$.

$\therefore$ Expected value of a random variable X is $E(X) = \sum_{i=1}^{n} x_i p_i$.

The above formula for finding the expected value of a random variable X is used only if X is a discrete random variable which takes the values $x_1, x_2, ..., x_n$ with probability mass function $p(x_i) = P[X = x_i], \ i = 1, 2, ..., n.$

But, if X is a continuous random variable having the probability density function $f(x)$, then in place of summation we will use integration and in this case, the expected value of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

The expectation, as defined above, agrees with the logical/theoretical argument also as is illustrated in the following example.

Suppose, a fair coin is tossed twice, then answer to the question, "How many heads do we expect theoretically/logically in two tosses?" is obviously 1 as the coin is unbiased and hence we will undoubtedly expect one head in two tosses. Expectation actually means "what we get on an average"? Now, let us obtain the expected value of the above question using the formula.

Let X be the number of heads in two tosses of the coin and we are to obtain E(X), i.e. expected number of heads. As X is the number of heads in two tosses of the coin, therefore X can take the values 0, 1, 2 and its probability distribution is given as

$X:$	0	1	2
$p(x):$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

[Refer Unit 5 of MST-003]

$$\therefore \ E(X) = \sum_{i=1}^{3} x_i p_i$$

$$= x_1 p_1 + x_2 p_2 + x_3 p_3$$

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

So, we get the same answer, i.e. 1 using the formula also.

So, expectation of a random variable is nothing but the average (mean) taken over all the possible values of the random variable or it is the value which we get on an average when a random experiment is performed repeatedly.

Remark 1: Sometimes summations and integrals as considered in the above definitions may not be convergent and hence expectations in such cases do not exist. But we will deal only those summations (series) and integrals which are convergent as the topic regarding checking the convergence of series or integrals is out of the scope of this course. You need not to bother as to whether the series or integral is convergent or not, i.e. as to whether the expectation exists or not as we are dealing with only those expectations which exist.

Example 1: If it rains, a rain coat dealer can earn Rs 500 per day. If it is a dry day, he can lose Rs 100 per day. What is his expectation, if the probability of rain is 0.4?

Solution: Let X be the amount earned on a day by the dealer. Therefore, X can take the values Rs 500, $-$ Rs 100 ($\because$ loss of Rs 100 is equivalent to negative of the earning of Rs100).

$\therefore$ Probability distribution of X is given as

	Rainy Day	Dry day
$X(\text{in Rs.})$:	500	-100
$p(x)$:	0.4	0.6

Hence, the expectation of the amount earned by him is

$$E(X)=\sum_{i=1}^{2} x_i p_i = x_1p_1 + x_2p_2$$

$$= (500)(0.4)+(-100)(0.6) = 200-60 = 140$$

Thus, his expectation is Rs 140, i.e. on an overage he earns Rs 140 per day.

Example 2: A player tosses two unbiased coins. He wins Rs 5 if 2 heads appear, Rs 2 if one head appears and Rs1 if no head appears. Find the expected value of the amount won by him.

Solution: In tossing two unbiased coins, the sample space, is

$S = \{HH, HT, TH, TT\}$.

$\therefore$ $P[2\,\text{heads}]=\frac{1}{4}$, $P(\text{one head})=\frac{2}{4}$, $P(\text{no head})=\frac{1}{4}$.

Let X be the amount in rupees won by him

$\therefore$ X can take the values 5, 2 and 1 with

$P[X=5] = P(2\text{heads}) = \frac{1}{4}$,

$P[X=2] = P[1\text{ Head}] = \frac{2}{4}$, and

$P[X=1] = P[\text{no Head}] = \frac{1}{4}$.

$\therefore$ Probability distribution of X is

X:	5	2	1
p(x)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Expected value of X is given as

$$E(X) = \sum_{i=1}^{3} x_i p_i = x_1p_1 + x_2p_2 + x_3p_3$$

$$= 5\left(\frac{1}{4}\right) + 2\left(\frac{2}{4}\right) + 1\left(\frac{1}{4}\right) = \frac{5}{4} + \frac{4}{4} + \frac{1}{4} = \frac{10}{4} = 2.5.$$

Thus, the expected value of amount won by him is Rs 2.5.

Example 3: Find the expectation of the number on an unbiased die when thrown.

Solution: Let X be a random variable representing the number on a die when thrown.

$\therefore$ X can take the values 1, 2, 3, 4, 5, 6 with

$$P[X=1] = P[X=2] = P[X=3] = P[X=4] = P[X=5] = P[X=6] = \frac{1}{6}.$$

Thus, the probability distribution of X is given by

X:	1	2	3	4	5	6
p(x):	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Hence, the expectation of number on the die when thrown is

$$E(X) = \sum_{i=1}^{6} x_i p_i = 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6} = \frac{21}{6} = \frac{7}{2}$$

Example 4: Two cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the expected value for the number of aces.

Solution: Let A_1, A_2 be the events of getting ace in first and second draws, respectively. Let X be the number of aces drawn. Thus, X can take the values 0, 1, 2 with

$$P[X=0] = P[\text{no ace}] = P\left[\bar{A}_1 \cap \bar{A}_2\right]$$

$$= P\left[\bar{A}_1\right]P\left[\bar{A}_2\right] \quad \left[\begin{array}{l}\because \text{cards are drawn with replacement} \\ \text{and hence the events are independent}\end{array}\right]$$

$$= \frac{48}{52} \times \frac{48}{52} = \frac{12}{13} \times \frac{12}{13} = \frac{144}{169},$$

$$P[X=1] = [\text{one Ace and one other card}]$$

$$= P\left[\left(A_1 \cap \bar{A}_2\right) \cup \left(\bar{A}_1 \cap A_2\right)\right]$$

$$= P\left[A_1 \cap \bar{A}_2\right] + P\left[\bar{A}_1 \cap A_2\right] \quad \left[\begin{array}{l}\text{By Addition theorem of probability} \\ \text{for mutually exclusive events}\end{array}\right]$$

$$= P[A_1]P\left[\bar{A}_2\right] + P\left[\bar{A}_1\right]P[A_2] \quad \left[\begin{array}{l}\text{By multiplication theorem of} \\ \text{probability for independent events}\end{array}\right]$$

$$= \frac{4}{52} \times \frac{48}{52} + \frac{48}{52} \times \frac{4}{52} = \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} = \frac{24}{169}, \text{ and}$$

$$P[X=2] = P[\text{both aces}] = P[A_1 \cap A_2]$$

$$= P[A_1]P[A_2] = \frac{4}{52} \times \frac{4}{52} = \frac{1}{169}.$$

Hence, the probability distribution of random variable X is

$X:$	0	1	2
$p(x):$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

$\therefore$ The expected value of X is given by

$$E(X) = \sum_{i=1}^{3} x_i p_i = 0 \times \frac{144}{169} + 1 \times \frac{24}{169} + 2 \times \frac{1}{169} = \frac{26}{169} = \frac{2}{13}$$

Example 5: For a continuous distribution whose probability density function is given by:

$$f(x) = \frac{3x}{4}(2-x),\ 0 \le x \le 2,\ \text{find the expected value of X.}$$

Solution: Expected value of a continuous random variable X is given by

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^2 x \frac{3x}{4}(2-x)\,dx = \frac{3}{4}\int_0^2 x^2(2-x)\,dx$$

$$= \frac{3}{4}\int_0^2 \left(2x^2 - x^3\right)dx = \frac{3}{4}\left[2\frac{x^3}{3} - \frac{x^4}{4}\right]_0^2 = \frac{3}{4}\left[2\frac{(2)^3}{3} - \frac{(2)^4}{4} - 0\right]$$

$$= \frac{3}{4}\left[\frac{16}{3} - \frac{16}{4}\right] = \frac{3}{4} \times \frac{16}{12} = 1$$

Now, you can try the following exercises.

E1) You toss a fair coin. If the outcome is head, you win Rs 100; if the outcome is tail, you win nothing. What is the expected amount won by you?

E2) A fair coin is tossed until a tail appears. What is the expectation of number of tosses?

E3) The distribution of a continuous random variable X is defined by

$$f(x) = \begin{cases} x^3 & , \ 0 < x \le 1 \\ (2-x)^3 & , \ 1 < x \le 2 \\ 0 & , \ \text{elsewhere} \end{cases}$$

Obtain the expected value of X.

Let us now discuss some properties of expectation in the next section.

8.3 PROPERTIES OF EXPECTATION OF ONE-DIMENSIONAL RANDOM VARIABLE

Properties of mathematical expectation of a random variable X are:

1. E(k) = k, where k is a constant
2. E(kX) = kE(X), k being a constant.
3. E(aX + b) = aE(X) + b, where a and b are constants

Proof:

Discrete case:

Let X be a discrete r.v. which takes the values $x_1, x_2, x_3, \ldots$ with respective probabilities $p_1, p_2, p_3, \ldots$

1. $E(k) = \sum_i k\, p_i$ [By definition of the expectation]

$= k \sum_i p_i$

$= k(1) = k$ $\left[\because \text{sum of probabilities of all the possible value of r.v. is 1}\right]$

2. $E(kX) = \sum_i (kx_i)\, p_i$ [By def.]

$= k \sum_i x_i\, p_i$

$= k\, E(X)$

3. $E(a\,X + b) = \sum_i (ax_i + b)\, p_i$ [By def.]

$= \sum_i (ax_i p_i + bp_i) = \sum_i ax_i p_i + \sum_i bp_i = a\sum_i x_i p_i + b\sum_i p_i$

$= aE(X) + b(1) \ = aE(X) + b$

Continuous Case:

Let X be continuous random variable having f(x) as its probability density function. Thus,

1. $E(k) = \int_{-\infty}^{\infty} kf(x)dx$ [By def.]

$= k\int_{-\infty}^{\infty} f(x)dx$

$= k(1) = k$ $\left[\begin{array}{l}\because \text{integral of the p.d.f. over} \\ \text{the entire range is 1}\end{array}\right]$

2. $E(kX) = \int_{-\infty}^{\infty} (kx)f(x)dx$ [By def.]

$= k\int_{-\infty}^{\infty} xf(x)dx = kE(X)$

3. $E(aX+b) = \int_{-\infty}^{\infty} (ax+b)f(x)dx = \int_{-\infty}^{\infty} (ax)f(x)dx + \int_{-\infty}^{\infty} bf(x)dx$

$= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx = aE(X)+b(1) = aE(X)+b$

Example 6: Given the following probability distribution:

X	–2	–1	0	1	2
$p(x)$	0.15	0.30	0	0.30	0.25

Find i) E(X)

ii) E(2X + 3)

iii) $E(X^2)$

iv) E(4X – 5)

Solution

i) $E(X) = \sum_{i=1}^{5} x_i p_i = x_1p_1 + x_2p_2 + x_3p_3 + x_4p_4 + x_5p_5$

$= (-2)(0.15) + (-1)(0.30) + (0)(0) + (1)(0.30) + (2)(0.25)$

$= -0.3 - 0.3 + 0 + 0.3 + 0.5 = 0.2$

ii) E(2X + 3) = $2E(X)+3$ [Using property 3 of this section]

= 2(0.2) + 3 [Using solution (i) of the question]

= 0.4 + 3 = 3.4

iii) $E(X^2) = \sum_{i=1}^{5} x_i^2 p_i$ [By def.]

$$= x_1^2p_1 + x_2^2p_2 + x_3^2p_3 + x_4^2p_4 + x_5^2p_5$$

$$= (-2)^2(0.15) + (-1)^2(0.30) + (0)^2(0) + (1)^2(0.30) + (2)^2(0.25)$$

$$= (4)(0.15) + (1)(0.30) + (0) + (1)(0.30) + (4)(0.25)$$

$$= 0.6 + 0.3 + 0 + 0.3 + 1 = 2.2$$

iv) $E(4X-5) = E[4X + (-5)]$

$$= 4E(X) + (-5) \qquad \text{[Using property 3 of this section]}$$

$$= 4(0.2) - 5$$

$$= 0.8 - 5 = -4.2$$

Here is an exercise for you.

E4) If X is a random variable with mean 'μ' and standard deviation 'σ', then what is the expectation of $Z = \frac{X-\mu}{\sigma}$?

[**Note:** Here Z so defined is called standard random variate.]

Let us now express the moments and other measures for a random variable in terms of expectations in the following section.

8.4 MOMENTS AND OTHER MEASURES IN TERMS OF EXPECTATIONS

Moments

The moments for frequency distribution have already been studied by you in Unit 3 of MST-002. Here, we deal with moments for probability distributions. The r^{th} order moment about any point 'A' (say) of variable X already defined in Unit 3 of MST-002 is given by:

$$\mu_r' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^r}{\sum_{i=1}^{n} f_i}$$

So, the r^{th} order moment about any point 'A' of a random variable X having probability mass function $P[X = x_i] = p(x_i) = p_i$ is defined as

$$\mu_r' = \frac{\sum_{i=1}^{n} p_i (x_i - A)^r}{\sum_{i=1}^{n} p_i}$$

[Replacing frequencies by probabilities as discussed in Sec. 8.2 of this unit.]

$$= \sum_{i=1}^{n} p_i (x_i - A)^r \qquad \left[\because \sum_{i=1}^{n} p_i = 1\right]$$

The above formula is valid if X is a discrete random variable. But, if X is a continuous random variable having probability density function f(x), then

r^{th} order moment about A is defined as $\mu_r' = \int_{-\infty}^{\infty} (x-A)^r f(x)dx.$

So, r^{th} order moment about any point 'A' of a random variable X is defined as

$$\mu_r' = \begin{cases} \sum_i p_i (x_i - A)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x-A)^r f(x)dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X - A)^r$$

Similarly, r^{th} order moment about mean (μ) i.e. r^{th} order central moment is defined as

$$\mu_r = \begin{cases} \sum_i p_i (x_i - \mu)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X-\mu)^r = E\left[X - E(X)\right]^r$$

Variance

Variance of a random variable X is second order central moment and is defined as

$$\mu_2 = V(X) = E[X-\mu]^2 = E\left[X - E(X)\right]^2$$

Also, we know that

$$V(X) = \mu_2' - (\mu_1')^2$$

where μ_1', μ_2' be the moments about origin.

$\therefore$ We have $V(X) = E(X^2) - \left[E(X)\right]^2$

$$\left[\because \mu_1' = E[X-0]^1 = E(X), \text{ and } \mu_2' = E[X-0]^2 = E(X^2)\right]$$

Theorem 8.1: If X is a random variable, then $V(aX+b) = a^2 V(X)$, where a and b are constants.

Proof: $V(aX+b) = E\left[(aX+b) - E(aX+b)\right]^2$ [By def. of variance]

$$= E\left[aX + b - \left(aE(X) + b\right)\right]^2 \quad \text{[Using property 3 of Sec. 8.3]}$$

$$= E\left[aX + b - aE(X) - b\right]^2$$

$$= E\left[a\left\{X - E(X)\right\}\right]^2$$

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

$$= a^2E\left[X - E\left(X\right)\right]^2 \qquad \text{[Using property 2 of section 8.3]}$$

$$= a^2V\left(X\right) \qquad \text{[By definition of Variance]}$$

Cor. (i) $V\left(aX\right) = a^2V\left(X\right)$

(ii) V(b) = 0

(iii) V(X + b) = V(X)

Proof: (i) This result is obtained on putting b = 0 in the above theorem.

(ii) This result is obtained on putting a = 0 in the above theorem.

(iii) This result is obtained on putting a = 1 in the above theorem.

Covariance

For a bivariate frequency distribution, you have already studied in Unit 6 of MST-002 that covariance between two variables X and Y is defined as

$$\text{Cov}\left(X, Y\right) = \frac{\sum_i f_i\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum_i f_i}$$

$\therefore$ For a bivariate probability distribution, Cov (X, Y) is defined as

$$\text{Cov(X, Y)} = \begin{cases} \sum_i p_{ij}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right), \text{ if } \left(X, Y\right) \text{ is two-dimensional discrete r.v.} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x - \bar{x}\right)\left(y - \bar{y}\right)f\left(x, y\right)dydx, \text{ if } \left(X, Y\right) \text{is two dimensional continuous r.v.} \end{cases}$$

where $p_{ij} = P\left[X = x_i, Y = y_j\right]$

$$= E\left(X - \bar{X}\right)\left(Y - \bar{Y}\right) \qquad \text{[By definition of expectation]}$$

$$= E\left[\left(X - E\left(X\right)\right)\left(Y - E\left(Y\right)\right)\right] \qquad \left[\begin{array}{l}\because E(X) = \text{Mean of X i.e. } \bar{X}, \\ E(Y) = \text{Mean of Y i.e. } \bar{Y}\end{array}\right]$$

On simplifying,

Cov(X, Y) = E(XY) – E(X) E(Y).

Now, if X and Y are independent random variables then, by multiplication theorem,

E(XY) = E(X)E(Y) and hence in this case Cov(X, Y) = 0.

Remark 2:

i) If X and Y are independent random variables, then

V(X + Y) = V(X) +V(Y).

Proof: $V(X+Y) = E\left[(X+Y) - E(X+Y)\right]^2$

$$= E\left[X+Y-E(X)-E(Y)\right]^2$$

$$= E\left[\{X-E(X)\}+\{Y-E(Y)\}\right]^2$$

$$= E\left[\{X-E(X)\}^2+\{Y-E(Y)\}^2+2\{X-E(X)\}\{Y-E(Y)\}\right]$$

$$= E\left[X-E(X)\right]^2+E\left[Y-E(Y)\right]^2+2E\left[(X-E(X))(Y-E(Y))\right]$$

$$= V(X)+V(Y)+2\text{Cov}(X,Y)$$

$$= V(X)+V(Y)+0 \qquad [\because \text{ X and Y are independent}]$$

$$= V(X)+V(Y)$$

ii) If X and Y are independent random variables, then

$V(X - Y) = V(X) + V(Y)$.

Proof: This can be proved in the similar manner as done is Remark 2(i) above.

iii) If X and Y are independent random variables, then

$V(aX+bY) = a^2V(X)+b^2V(Y)$.

Proof: Prove this result yourself proceeding in the similar fashion as in proof of Remark 2(i).

Mean Deviation about Mean

Mean deviation about mean in context of frequency distribution is

$$\frac{\sum_{i=1}^{n} f_i \,|\, x_i - \bar{x} \,|}{\sum_{i=1}^{n} f_i}, \text{ and}$$

therefore, mean deviation about mean in context of probability distribution is

$$\frac{\sum_{i=1}^{n} p_i \,|\, x - \text{mean} \,|}{\sum_{i=1}^{n} p_i} = \sum_{i=1}^{n} p_i \,|\, x - \text{mean} \,|$$

$\therefore$ by definition of expectation, we have

M.D. about mean = $E|X - \text{Mean}|$

$= E|X - E(X)|$

$$= \begin{cases} \sum p_i \,|\, x - \text{Mean} \,| & \text{for discrete r.v} \\ \int_{-\infty}^{\infty} |\, x - \text{Mean} \,|\, f(x)\,dx & \text{for continuous r.v} \end{cases}$$

Note: Other measures as defined for frequency distributions in MST-002 can be defined for probability distributions also and hence can be expressed in terms of the expectations in the manner as the moments; variance and covariance have been defined in this section of the Unit.

Example 7: Considering the probability distribution given in Example 6, obtain

i) V(X)

ii) V(2X + 3).

Solution:

(i) $V(X) = E(X^2) - [E(X)]^2$

$= 2.2 - (0.2)^2$ $\left[\begin{array}{l}\text{The values have already been obtained} \\ \text{in the solution of Example 6}\end{array}\right]$

$= 2.2 - 0.04 = 2.16$

(ii) $V(2X+3) = (2)^2\, V(X)$ [Using the result of Theorem 8.1]

$= 4V(X) = 4(2.16) = 8.64$

Example 8: If X and Y are independent random variables with variances 2 and 3 respectively, find the variance of 3X + 4Y.

Solution: $V(3X + 4Y) = (3)^2\, V(X) + (4)^2\, V(Y)$ [By Remark 3 of Section 8.4]

$= 9(2) + 16(3) = 18 + 48 = 66$

Here are two exercises for you:

E5) If X is a random variable with mean μ and standard deviation σ, then find the variance of standard random variable $Z = \dfrac{X - \mu}{\sigma}$.

E6) Suppose that X is a random variable for which E(X) = 10 and V(X) = 25. Find the positive values of a and b such that Y = aX – b has expectation 0 and variance 1.

8.5 ADDITION AND MULTIPLICATION THEOREMS OF EXPECTATION

Now, we are going to deal with the properties of expectation in case of two-dimensional random variable. Two important properties, i.e. addition and multiplication laws of expectation are discussed in the present section.

Addition Theorem of Expectation

Theorem 8.2: If X and Y are random variables, then $E(X+Y)=E(X)+E(Y)$

Proof:

Discrete case:

Let (X, Y) be a discrete two-dimensional random variable which takes up the values (x_i, y_j) with the joint probability mass function $p_{ij} = P\left[X = x_i \cap Y = y_j\right]$.

Then, the probability distribution of X is given by

$$p_i = p(x_i) = P[X = x_i]$$

$$= P[X = x_i \cap Y = y_1] + P[X = x_i \cap Y = y_2] + \ldots \begin{bmatrix} \because \text{event } X = x_i \text{ can happen with} \\ Y{=}y_1 \text{ or } Y{=}y_2 \text{ or } Y{=}y_3 \text{ or} \ldots \end{bmatrix}$$

$$= p_{i1} + p_{i2} + p_{i3} + \ldots$$

$$= \sum_j p_{ij}$$

Similarly, the probability distribution of Y is given by

$$p_j' = p(y_j) = P\left[Y = y_j\right] = \sum_i p_{ij}$$

$$\therefore \; E(X) = \sum_i x_i p_i, \; E(Y) = \sum_j y_j p_j' \text{ and } E(X+Y) = \sum_i \sum_j (x_i + y_j) p_{ij}$$

Now $E(X+Y) = \sum_i \sum_j (x_i + y_j) p_{ij}$

$$= \sum_i \sum_j x_i p_{ij} + \sum_i \sum_j y_j p_{ij}$$

$$= \sum_i x_i \sum_j p_{ij} + \sum_j y_j \sum_i p_{ij}$$

[$\because$ in the first term of the right hand side, x_i is free from j and hence can be taken outside the summation over j; and in second term of the right hand side, y_j is free from i and hence can be taken outside the summation over i.]

$$\therefore \; E(X+Y) = \sum_i x_i p_i + \sum_j y_j p_j' = E(X) + E(Y)$$

Continuous Case:

Let (X, Y) be a bivariate continuous random variable with probability density function $f(x, y)$. Let $f(x)$ and $f(y)$ be the marginal probability density functions of random variables X and Y respectively.

$\therefore\ E(X)=\int_{-\infty}^{\infty} x\ f(x)dx,\ E(Y)=\int_{-\infty}^{\infty} y\ f(y)dy,$

and $E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dy\,dx$.

Now, $E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dy\,dx$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,f(x,y)dy\,dx+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\,f(x,y)dy\,dx$$

$$=\int_{-\infty}^{\infty}x\left(\int_{-\infty}^{\infty}f(x,y)dy\right)dx+\int_{-\infty}^{\infty}y\left(\int_{-\infty}^{\infty}f(x,y)dx\right)dy$$

[$\because$ in the first term of R.H.S., x is free from the integral w.r.t. y and hence can be taken outside this integral. Similarly, in the second term of R.H.S, y is free from the integral w.r.t. x and hence can be taken outside this integral.]

$$=\int_{-\infty}^{\infty}x\,f(x)dx+\int_{-\infty}^{\infty}y\,f(y)dy\ \left[\begin{array}{l}\text{Refer to the definition of marginal density}\\ \text{function given in Unit 7 of this course}\end{array}\right]$$

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

Remark 3: The result can be similarly extended for more than two random variables.

Multiplication Theorem of Expectation

Theorem 8.3: If X and Y are independent random variables, then

E(XY) = E(X) E(Y)

Proof:

Discrete Case:

Let (X, Y) be a two-dimensional discrete random variable which takes up the values (x_i, y_j) with the joint probability mass function

$p_{ij}=P\left[X=x_i\cap Y=y_j\right]$. Let p_i and $p_j^{'}$ be the marginal probability mass functions of X and Y respectively.

$\therefore\ E(X)=\sum_i x_i p_i,\ E(Y)=\sum_j y_j p_j^{'},$ and

$$E(XY)=\sum_i\sum_j(x_i y_j)p_{ij}$$

But as X and Y are independent,

$\therefore\ p_{ij}=P\left[X=x_i\cap Y=y_j\right]$

$$= P[X = x_i]P[Y = y_j] \quad \left[\begin{array}{l}\because \text{ if events A and B are independent,} \\ \text{then } P(A \cap B) = P(A)P(B)\end{array}\right]$$

$$= p_i p_j'$$

Hence, $E(XY) = \sum_i \sum_j (x_i y_j) p_i p_j'$

$$= \sum_i \sum_j x_i y_j p_i p_j'$$

$$= \sum_i \sum_j (x_i p_i y_j p_j')$$

$$= \sum_i x_i p_i \sum_j y_j p_j' \quad \left[\begin{array}{l}\because\ x_i p_i \text{ is free from j and hence can be} \\ \text{taken outside the summation over j}\end{array}\right]$$

$$= E(X)\, E(Y)$$

Continuous Case:

Let (X, Y) be a bivariate continuous random variable with probability density function f(x, y). Let f(x) and f(y) be the marginal probability density function of random variables X and Y respectively.

$$\therefore\ E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx,\ E(Y) = \int_{-\infty}^{\infty} y\, f(y)\, dy,$$

and $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x, y)\, dy\, dx$.

Now $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x, y)\, dy\, dx$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x) f(y)\, dy\, dx \quad \left[\begin{array}{l}\because \text{ X and Y are independent, } f(x,y) = f(x)f(y) \\ \text{(see Unit 7 of this course)}\end{array}\right]$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (x\, f(x))(y f(y)) \right) dy\, dx$$

$$= \left(\int_{-\infty}^{\infty} x\, f(x)\, dx \right) \left(\int_{-\infty}^{\infty} y\, f(y)\, dy \right)$$

$$= E(X)E(Y)$$

Remark 4: The result can be similarly extended for more than two random variables.

Example 8: Two unbiased dice are thrown. Find the expected value of the sum of number of points on them.

Solution: Let X be the number obtained on the first die and Y be the number obtained on the second die, then

$E(X) = \frac{7}{2}$ and $E(Y) = \frac{7}{2}$ [See Example 3 given in Section 8.2]

$\therefore$ The required expected value $= E(X + Y)$

$$= E(X) + E(Y) \quad \left[\begin{array}{l}\text{Using addition theorem} \\ \text{of expectation}\end{array}\right]$$

$$= \frac{7}{2} + \frac{7}{2} = 7$$

Remark 5: This example can also be done considering one random variable only as follows:

Let X be the random variable denoting "the sum of numbers of points on the dice", then the probability distribution in this case is

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

and hence $E(X) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + ... + 12 \times \frac{1}{36} = 7$

Example 9: Two cards are drawn one by one with replacement from 8 cards numbered from 1 to 8. Find the expectation of the product of the numbers on the drawn cards.

Solution: Let X be the number on the first card and Y be the number on the second card. Then probability distribution of X is

X	1	2	3	4	5	6	7	8
p(x)	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

and the probability distribution of Y is

Y	1	2	3	4	5	6	7	8
p(y)	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$\therefore \quad E(X) = E(Y) = 1 \times \frac{1}{8} + 2 \times \frac{1}{8} + ... + 8 \times \frac{1}{8}$$

$$= \frac{1}{8}(1+2+3+4+5+6+7+8) = \frac{1}{8}(36) = \frac{9}{2}$$

Thus, the required expected value is

$E(XY) = E(X)E(Y)$ [Using multiplication theorem of expectation]

$$= \frac{9}{2} \times \frac{9}{2} = \frac{81}{4}.$$

Expectation of Linear Combination of Random Variables

Theorem 8.4: Let $X_1, X_2, ..., X_n$ be any n random variables and if $a_1, a_2, ..., a_n$ are any n constants, then

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n)$$

[**Note :** Here $a_1X_1 + a_2X_2 + ... + a_nX_n$ is a linear combination of $X_1, X_2, ..., X_n$]

Proof: Using the addition theorem of expectation, we have

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n) = E(a_1X_1) + E(a_2X_2) + ... + E(a_nX_n)$$

$$= a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n).$$

[Using second property of Section 8.3 of the unit]

Now, you can try the following exercises.

E7) Two cards are drawn one by one with replacement from ten cards numbered 1 to 10. Find the expectation of the sum of points on two cards.

E8) Find the expectation of the product of number of points on two dice.

Now before ending this unit, let's summarize what we have covered in it.

8.6 SUMMARY

The following main points have been covered in this unit:

1) Expected value of a random variable X is defined as

$$E(X) = \sum_{i=1}^{n} x_i p_i \text{ , if X is a discrete random variable}$$

$$= \int_{-\infty}^{\infty} x f(x)\, dx \text{ , if X is a continuous random variable.}$$

2) Important properties of expectation are:

   i) E(k) = k, where k is a constant.

   ii) E(kX) = kE(X), k being a constant.

   iii) E(aX + b) = aE(X) + b, where a and b are constants

   iv) Addition theorem of Expectation is stated as:

   If X and Y are random variables, then $E(X+Y) = E(X) + E(Y)$.

   v) Multiplication theorem of Expectation is stated as:
   If X and Y are independent random variables, then
   E(XY) = E(X)E(Y).

vi) If $X_1, X_2, ..., X_n$ be any n random variables and if $a_1, a_2, ..., a_n$ are any n constants, then

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n).$$

3) Moments and other measures in terms of expectation are given as:

i) r^{th} order moment about any point is given as

$$\mu_r' = \begin{cases} \sum_i p_i (x_i - A)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x - A)^r f(x) dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X - A)^r$$

ii) Variance of a random variable X is given as

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

iii) $$\text{Cov(X,Y)} = \begin{cases} \sum_i p_i (x_i - \bar{x})(y_i - \bar{y}), & \text{if } (X, Y) \text{ is discrete r.v.} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \bar{x})(y - \bar{y}) f(x, y) dydx, & \text{if } (X, Y) \text{ is continuous r.v.} \end{cases}$$

$$= E[(X - E(X))(Y - E(Y))]$$

$$= E(XY) - E(X)\,E(Y).$$

iv) M.D. about mean $= E|X - E(X)|$

$$= \begin{cases} \sum p_i \,|\, x - \text{Mean} \,| & \text{for discrete r.v} \\ \int_{-\infty}^{\infty} |\, x - \text{Mean} \,| f(x) dx & \text{for continuous r.v} \end{cases}$$

If you want to see what our solutions to the exercises in the unit are, we have given them in the following section.

8.7 SOLUTIONS/ANSWERS

E1) Let X be the amount (in rupees) won by you.

$\therefore$ X can take the values 100, 0 with P[X = 100] = P[Head] = $\frac{1}{2}$, and

P[X = 0] = P[Tail] = $\frac{1}{2}$.

$\therefore$ probability distribution of X is

X :	100	0
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

and hence the expected amount won by you is

$$E(X) = 100 \times \frac{1}{2} + 0 \times \frac{1}{2} = 50.$$

E2) Let X be the number of tosses till tail turns up.

$\therefore$ X can take values 1, 2, 3, 4… with

P[X = 1] = P[Tail in the first toss] = $\frac{1}{2}$

P[X = 2] = P[Head in the first and tail in the second toss] = $\frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^2$,

P[X = 3] = P[HHT] = $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^3$, and so on.

$\therefore$ Probability distribution of X is

X:	1	2	3	4	5...
$p(x)$	$\frac{1}{2}$	$\left(\frac{1}{2}\right)^2$	$\left(\frac{1}{2}\right)^3$	$\left(\frac{1}{2}\right)^4$	$\left(\frac{1}{2}\right)^5$ ...

and hence

$$E(X) = 1 \times \frac{1}{2} + 2 \times \left(\frac{1}{2}\right)^2 + 3 \times \left(\frac{1}{2}\right)^3 + 4 \times \left(\frac{1}{2}\right)^4 + \ldots \qquad \ldots (1)$$

Multiplying both sides by $\frac{1}{2}$, we get

$$\frac{1}{2}E(X) = \left(\frac{1}{2}\right)^2 + 2 \times \left(\frac{1}{2}\right)^3 + 3 \times \left(\frac{1}{2}\right)^4 + 4 \times \left(\frac{1}{2}\right)^5 + \ldots$$

$$\Rightarrow \frac{1}{2}E(X) = \left(\frac{1}{2}\right)^2 + 2 \times \left(\frac{1}{2}\right)^3 + 3 \times \left(\frac{1}{2}\right)^4 + \ldots \qquad \ldots (2)$$

[Shifting the position one step towards right so that we get the terms having same power at the same positions as that in (1)]

Now, subtracting (2) from (1), we have

$$E(X) - \frac{1}{2}E(X) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \ldots$$

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

$$\Rightarrow E(X) = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \ldots$$

(Which is an infinite G.P. with first term a = 1 and common ratio $r = \frac{1}{2}$)

$$= \frac{1}{1-\frac{1}{2}} \quad [\because S_\infty = \frac{a}{1-r} \text{ (see Unit 3 of course MST - 001)}]$$

$$= \frac{1}{\frac{1}{2}} = 2.$$

E3) $E(X) = \int_{-\infty}^{\infty} x f(x) dx$

$$= \int_{-\infty}^{0} x f(x) dx + \int_{0}^{1} x f(x) dx + \int_{1}^{2} x f(x) dx + \int_{2}^{\infty} x f(x) dx$$

$$= \int_{-\infty}^{0} x(0) dx + \int_{0}^{1} x(x^3) dx + \int_{1}^{2} x(2-x)^3 dx + \int_{2}^{\infty} x(0) dx$$

$$= 0 + \int_{0}^{1} x^4 dx + \int_{1}^{2} x\left[8 - x^3 - 6x(2-x)\right] dx + 0$$

$$= \int_{0}^{1} x^4 dx + \int_{1}^{2} \left(8x - x^4 - 12x^2 + 6x^3\right) dx$$

$$= \left[\frac{x^5}{5}\right]_0^1 + \left[8\frac{x^2}{2} - \frac{x^5}{5} - 12\frac{x^3}{3} + 6\frac{x^4}{4}\right]_1^2$$

$$= \frac{1}{5} + \left[\left\{\frac{8(2)^2}{2} - \frac{(2)^5}{5} - \frac{12(2)^3}{3} + \frac{6(2)^4}{4}\right\} - \left\{\frac{8(1)^2}{2} - \frac{(1)^5}{5} - \frac{12(1)^3}{3} + \frac{6(1)^4}{4}\right\}\right]$$

$$= \frac{1}{5} + \left[\left\{16 - \frac{32}{5} - 32 + 24\right\} - \left\{4 - \frac{1}{5} - 4 + \frac{3}{2}\right\}\right]$$

$$= \frac{1}{5} + \left[\frac{8}{5} - \frac{13}{10}\right] = \frac{1}{5} + \frac{3}{10} = \frac{1}{2}.$$

E4) As X is a random variable with mean μ,

$\therefore$ E(X) = μ ... (1)

$\left[\begin{array}{l}\because \text{expectation is nothing but simply the average taken over all} \\ \text{the possible values of random variable as defined in Sec. 8.2}\end{array}\right]$

Now, $E(Z) = E\left(\frac{X-\mu}{\sigma}\right)$

$$= E\left[\frac{1}{\sigma}(X-\mu)\right]$$

$$= \frac{1}{\sigma}E[X-\mu] \qquad \text{[Using Property 2 of Sec. 8.3]}$$

$$= \frac{1}{\sigma}[E(X)-\mu] \qquad \text{[Using Property 3 of Sec. 8.3]}$$

$$= \frac{1}{\sigma}[\mu-\mu] \qquad \text{[Using (1)]}$$

$$= 0$$

Note: Mean of standard random variable is zero.

E5) Variance of standard random variable $Z = \frac{X-\mu}{\sigma}$ is given as

$$V(Z) = V\left(\frac{X-\mu}{\sigma}\right) = V\left(\frac{X}{\sigma}-\frac{\mu}{\sigma}\right)$$

$$= V\left[\frac{1}{\sigma}X + \left(-\frac{\mu}{\sigma}\right)\right]$$

$$= \left(\frac{1}{\sigma}\right)^2 V(X) \qquad \left[\begin{array}{l}\text{Using the result of the Theorem 8.1}\\ \text{of Sec. 8.5 of this unit}\end{array}\right]$$

$$= \frac{1}{\sigma^2}V(X)$$

$$= \frac{1}{\sigma^2}(\sigma^2) = 1 \qquad \left[\begin{array}{l}\because \text{it is given that the standard deviation}\\ \text{of X is and hence its variance is } \sigma^2\end{array}\right]$$

Note: The mean of standard random variate is '0' [See **(E4)**] and its variance is 1.

E6) Given that $E(Y) = 0 \Rightarrow E(aX - b) = 0 \Rightarrow a\,E(X) - b = 0$

$\Rightarrow a(10) - b = 0$

$\Rightarrow 10a - b = 0$ ... (1)

Also as $V(Y) = 1$,

hence $V(aX - b) = 1$

$\Rightarrow a^2V(X) = 1 \Rightarrow a^2(25) = 1 \Rightarrow a^2 = \frac{1}{25}$

$\Rightarrow a = \frac{1}{5}$ [$\because$ a is positive]

$\therefore$ From (1), we have

$$10\left(\frac{1}{5}\right) - b = 0 \Rightarrow 2 - b = 0 \Rightarrow b = 2$$

Hence, $a = \frac{1}{5}$, b = 2.

E7) Let X be the number on the first card and Y be the number on the second card. Then probability distribution of X is:

X	1	2	3	4	5	6	7	8	9	10
$p(x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

and the probability distribution of Y is

X	1	2	3	4	5	6	7	8	9	10
$p(x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

$$\therefore\ E(X) = E(Y) = 1\times\frac{1}{10}+2\times\frac{1}{10}+\ldots+10\times\frac{1}{10}$$

$$=\frac{1}{10}\left[1+2+3+4+5+6+7+8+9+10\right] = \frac{1}{10}(55) = 5.5$$

and hence the required expected value is

$$E(X+Y) = E(X)+E(Y) = 5.5 + 5.5 = 11$$

E8) Let X be the number obtained on the first die and Y be the number obtained on the second die.

Then $E(X) = E(Y) = \frac{7}{2}$. [See Example 3 given in Section 8.2]

Hence, the required expected value is

$$E(XY) = E(X)E(Y) \quad \text{[Using multiplication theorem of expectation]}$$

$$=\frac{7}{2}\times\frac{7}{2} = \frac{49}{4}.$$