2 Random Variable, Probability Mass Function, Probability Density Function, Distribution Function

2.1 INTRODUCTION :

In this chapter we shall discuss the distribution of probability by means of random variable. We are interested to find a numerical valued function of the sample points of a sample space. Suppose two coins are tossed. Then the sample space

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

Let X referes to the occurrence of head, then we have

$$P(X = 0) = P(\text{event } TT) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$$P(X = 1) = P(\text{event } HT) + P(\text{event } TH)$$

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

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

Also, $P(X = 0) + P(X = 1) + P(X = 2) = 1$

This is one kind of probability distribution and then the behaviour of it will be studid in the form of mathematical expection.

2.2 RANDOM VARIABLE :

A **random variable** is a real valued function defined on a sample space S. That is a random variable means a real number "X" associated with the outcomes of a random experiment.

e.g. suppose three coins are tossed and if X is the number of heads then X is a random variable which can assume the values 0, 1, 2, 3.

If a random variable assumes integer values such as 0, 1, 2, 3, ... then it is known as **discrete random variable.** e.g. Number of heads when we toss two coins, number of fans in a room, number of accidents taking place on a busy road.

If a random variable takes on all values within a certain interval, then it is called **continuous random variable.** e.g. height of the students, price of the commodity.

2.3 PROBABILITY DISTRIBUTION OF A RANDOM VARIABLE :

Probability distribution is a distribution of probability corresponding to different values of a random variable. It is quite different from frequency distribution as it can be constructed from a set of data obtained from an experiment or by observing a phenomenon. Probability distribution is of two types.

(i) Discrete Probability Distribution :

For the discrete random variable, the corresponding probability distribution is called **discrete distribution.** For example in tossing of 3 coins an X is the appearance of head then the discrete distribution is

X :	0	1	2	3
P :	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Note that this is analogous to that of frequency distribution with the porbabilities replacing relative frequencies.

(ii) Continuous Probability Distribution :

The distribution of probabilities corresponding to continuous random variable is called **coutinuous distribution.** For example for the height distribution of students.

X	50–55	55–60	60–65	65–70
P	$\frac{3}{20}$	$\frac{7}{20}$	$\frac{4}{20}$	$\frac{6}{20}$

(iii) Probability Function :

(a) Probability Mass Function :

Probability distribution function or **probability mass functions (pmf)** are those functions which give us probability corresponding to discrete random variable. Thus if X is a discrete random variable then $f(x) = P(X = x)$.

That is if $x_1, x_2, ..., x_k$ are the values of discrete random variable X with respective probabilities $p_1, p_2, ..., p_k$ where $p_1 + p_2 + ... + p_k = 1$ then probability mass function is given by $p_k = P(X = X_k) = f(x_k)$

The probability mass function is often the primary means of defining a discrete probability distribution and such functions exist for either single or multivariate random variables, given that the distribution is discrete.

The **characteristics** of a probability mass function are :

(i) $f(x) \geq 0$ for all x

(ii) $\sum_x f(x) = 1$

(iii) $P(X = x_i \text{ or } x_j) = P(X = x_i) + P(X = x_j)$

(b) Probability Density Function :

The probability function of a continuous random variable is called a **probability density function (pdf)** or frequency function.

The range of possible values is uncountably infinite for continuous random variables. So in this case, the distribution is defined by the probability density function $f(x)$ for the given range of random variable X.

The probability density function $f(x)$ is a function which, when integrated between a and b, gives the probability that the random variable will assume a value between a and b. That is,

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

Here we will talk about the probability in an interval, not at a particular point. For example

$$f(x) = 1 \quad , 0 < x < 2$$
$$= 0 \quad , \text{elsewhere}$$

If $f(x)\, dx$ is the probability that a random variable X takes the value in the small interval of width dx, $(x, x + dx)$ then $f(x)$ is called the probability density function.

The **characteristics** of pdf are :

(i) $f(x) \geq 0$ for all x

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

2.4 CHARACTERISTICS OF PROBABILITY DISTRIBUTION :

Distribution function : For discrete random variable the probability

$$F(x) = P(X \leq x) = \sum_{X \leq x} f(x)$$

is called the **cumulative distribution function** or simply **distribution function** of the random variable e.g. for the tossing of two coins, we have

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

For continuous random variable the **distribution function** is defined as

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

e.g. for the function $f(x) = 1 \quad , 0 < x < 2$

$= 0$, elsewhere,

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

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

Properties :

(i) Discrete r.v.

$P(a < X < b) = F(b) - F(a)$

$P(a \leq X < b) = P(X = a) + \{F(b) - F(a)\}$

$P(a < X \leq b) = \{F(b) - F(a)\} - P(X = b)$

$P(a \leq X \leq b) = \{F(b) - F(a)\}$
$+ P(X = a) - P(X = b)$

(ii) Continuous r.v.

$P(a < X < b) = P(a \le X < b) = P(a < X \le b)$

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

= Area between the curve $y = f(x)$, X–axis, the lines $x = a$, $x = b$

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0,$$

$$F(\infty) = \lim_{x \to \infty} F(x) = 1$$

$F(x) \le F(y)$ When $x < y$

For continuous case $F'(x) = f(x) \ge 0 \Rightarrow F(x)$ is an increasing function.

Illustration 1 : Which of the following define probability distributions :

(a) $f(x) = \dfrac{x-2}{10}$, $x = 0, 1, 2, 3, 4$

(b) $f(x) = \dfrac{x}{6}$, $x = 0, 1, 2, 3$

(c) $f(x) = 0$ **for** $x \le 0$

$= 8xe^{-4x^2}$ **for** $x > 0$

Solution :

(a) Here $f(0) = -\dfrac{2}{10}$, $f(1) = -\dfrac{1}{10}$

$\therefore$ $f(x) < 0 \Rightarrow f(x)$ is not probability distribution.

(b) Here $f(x) \ge 0$, for all x.

Also $\sum_x f(x) = f(0) + f(1) + f(2) + f(3)$

$$= 0 + \frac{1}{6} + \frac{2}{6} + \frac{3}{6} = 1$$

$\therefore$ $f(x)$ is a probability distribution.

(c) Here $f(x) \ge 0 \ \forall x$

$$\text{Also, } \int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{\infty} 8xe^{-4x^2}\,dx$$

$$= 0 + \left[\frac{e^{-4x^2}}{-1}\right]_0^{\infty} = 1$$

$\therefore$ $f(x)$ is a probability distribution.

Illustration 2 : When a dice is tossed, obtain the probability distribution of fours in two tosses.

Solution : Let X = number of fours in tossing a dice.

$$P(\text{getting 4 in a toss}) = \frac{1}{6},$$

$$P(\text{getting no 4 in a toss}) = \frac{5}{6}$$

For $X = 0$; $f(0)$ = P(not four and not four)

$$= \frac{5}{6} \times \frac{5}{6} = \frac{25}{36}$$

For $X = 1$; $f(1)$ = (four and not four) + P(not four and four)

$$= \frac{1}{6} \times \frac{5}{6} + \frac{5}{6} \times \frac{1}{6} = \frac{10}{36}$$

For $X = 2$; $f(2)$ = P(four and four)

$$= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

$\therefore$ The probability distribution is :

X	0	1	2
$f(x)$	$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$

Illustration 3 : A fair coin is tossed until a head or five tails occur. Obtain the probability distribution.

Solution : Let X = number of tails.

X	Sample points	Probability
0	H	$\frac{1}{2}$
1	TH	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
2	TTH	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$
3	TTTH	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$
4	TTTTH	$\frac{1}{32}$
5	TTTTT	$\frac{1}{32}$

Illustration 4 : For the following distribution find the value of k.

x	1	2	3	4	5
$f(x)$	0.1	k	.2	$3k$	.3

Solution : Since $f(x) \geq 0, \quad k \geq 0$

$\Sigma f(x) = 1 \Rightarrow 0.1 + k + 0.2 + 3k + 0.3 = 1$

$\Rightarrow 4k + 0.6 = 1$

$\Rightarrow 4k = 0.4 \Rightarrow k = 0.1$

Illustration 5 : For the probability function $f(x) = \dfrac{k}{1+x^2}, -\infty < x < \infty$, find k. (Nov. 2017)

Solution : We have $f(x) \geq 0 \Rightarrow k \geq 0$

Also $\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_{-\infty}^{\infty} \frac{k}{1+x^2}\,dx = 1$

$$\Rightarrow k\left[\tan^{-1} x\right]_{-\infty}^{\infty} = 1$$

$$\Rightarrow k\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1 \Rightarrow k = \frac{1}{4}$$

Illustration 6 : Is $f(x) = \dfrac{x}{6}$, $x = 0, 1, 2, 3, 4$ define probability distribution ? Justify your answer. (May 2017)

Solution : Here $f(x) \geq 0$ for all x.

Now, $\sum_x f(x) = f(0) + f(1) + f(2) + f(3) + f(4)$

$$= 0 + \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} = \frac{5}{3} > 1$$

$\therefore \quad f(x)$ is not a probability distribution.

2.5 TWO DIMENSIONAL RANDOM VARIABLES AND THEIR DISTRIBUTION FUNCTIONS :

Sometimes based on random experiment there exist two random variables, say X and Y, then we need to determine the relationship between them. Thus we may define **two dimensional random variable** as a pair (X, Y).

2.5.1 Discrete Random Variables :

Let X and Y be two discrete random variables. Let X assumes the value x and Y assume the value y. Then the probability for the intersection of two variables X and Y is given by

$$P(X = x, Y = y) = f(x, y)$$

It is also called as joint probability mass function.

Also $\sum_{\forall x}\sum_{\forall y} f(x, y) = 1$

2.5.2 Marginal Probability Function :

(a) For each fixed value of x (i.e. for individual random variable), the marginal probability distribution or **marginal probability function of X** is defined as

$$P(X = x) = f_1(x) = \sum_{\forall y} f(x, y)$$

where the summation run over all possible values of y.

Again, we have $\sum_{\forall x} f_1(x) = 1$.

(b) For each fixed value of y, the **marginal probability function of Y** is defined as

$$P(Y = y) = f_2(y) = \sum_{\forall x} f(x, y)$$

where the summation run over all possible values of x.

Again, we have $\sum_{\forall y} f(x, y) = 1$.

Now, we summarize above discussion in a table.

Let X has the values x_1, x_2, x_3, x_4 and Y has the values y_1, y_2, y_3.

Y \ X	x_1	x_2	x_3	x_4	Total $f_2(y)$
y_1	$f(x_1, y_1)$	$f(x_2, y_1)$	$f(x_3, y_1)$	$f(x_4, y_1)$	$f_2(y_1)$
y_2	$f(x_1, y_2)$	$f(x_2, y_2)$	$f(x_3, y_2)$	$f(x_4, y_2)$	$f_2(y_2)$
y_3	$f(x_1, y_3)$	$f(x_2, y_3)$	$f(x_3, y_3)$	$f(x_4, y_3)$	$f_3(y_3)$
Total $f_1(x)$	$f_1(x_1)$	$f_1(x_2)$	$f_1(x_3)$	$f_1(x_4)$	1.0

The total of all $f(x, y)$ values is equal to 1.

2.5.3 Independent Random Variables :

Let X and Y be two random variables then the **conditional probability distribution** (function) of X $= x$ given Y $= y$ (pre–fixed) is defined as

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} \text{ for all } x \text{ provided}$$

$f_2(y) \neq 0.$

It is dependent on y.

If $f_1(x \mid y) = f_1(x)$ for all x and y then conditional probability function is independent of y. That is when $f(x, y) = f_1(x) \cdot f_2(y)$ for all x and y.

The variables X and Y are called **independent random variables.**

Similary, the condition probability distribution function of Y $= y$ given X $= x$ is defined as

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} \text{ for all } y \text{ provided}$$

$f_1(x) \neq 0.$

Here, again if $f_2(y \mid x) = f_2(y)$ for all x and y then X and Y are called as **independent random variables.**

Illustration 1 : Let the discrete random variables X and Y have the joint probability distribution $f(x, y)$ is

Y \ X	1	2	3
1	0.1	0.4	0.1
2	0.2	0.2	0

Find (a) P(X + Y > 2) (b) P(X + Y < 4) (c) Marginal probability distribution of X and Y (d) The marginal probability distribution of Y on X (e) The conditional probability distribution of X given Y = 1 (*f*) The conditional probability of Y given X = 2.

Solution :

(a) We have X + Y > 2

$$\therefore \quad P(X + Y > 2) = f(1, 2) + f(2, 1) + f(2, 2) + f(3, 1) + f(3, 2)$$

$$= 0.2 + 0.4 + 0.2 + 0.1 + 0 = 0.9$$

(b) We have X + Y < 4

$$\therefore \quad P(X + Y < 4) = f(1, 1) + f(1, 2) + f(2, 1)$$

$$= 0.1 + 0.2 + 0.4 = 0.7$$

(c) We have the marginal probability distribution of X on Y are given by

$$f_1(1) = P(X = 1) = f(1, 1) + f(1, 2) = 0.1 + 0.2 = 0.3$$

$$f_1(2) = P(X = 2) = f(2, 1) + f(2, 2) = 0.4 + 0.2 = 0.6$$

$$f_1(3) = P(X = 3) = f(3, 1) + f(3, 2) = 0.1 + 0 = 0.1$$

(d) The marginal probability distribution of Y on X are given by

$$f_2(1) = P(Y = 1) = f(1, 1) + f(2, 1) + f(3, 1) = 0.1 + 0.4 + 0.1 = 0.6$$

$$f_2(2) = P(Y = 2) = f(1, 2) + f(2, 2) + f(3, 2) = 0.2 + 0.2 + 0 = 0.4$$

(e) The conditional probability distribution of X given Y = 1 is

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)}$$

For $x = 1$, $f_1(1 \mid 1) = \dfrac{f(1, 1)}{f_2(1)} = \dfrac{0.1}{0.6} = \dfrac{1}{6}$

For $x = 2$, $f_1(2 \mid 1) = \dfrac{f(2, 1)}{f_2(1)} = \dfrac{0.4}{0.6} = \dfrac{4}{6} = \dfrac{2}{3}$

For $x = 3$, $f_1(3 \mid 1) = \dfrac{f(3, 1)}{f_2(1)} = \dfrac{0.1}{0.6} = \dfrac{1}{6}$

(f) The conditional probability of Y given X = 2 is

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)}$$

For $y = 1$, $f_2(1 \mid 2) = \dfrac{f(2, 1)}{f_1(2)} = \dfrac{0.4}{0.6} = \dfrac{4}{6} = \dfrac{2}{3}$

For $y = 2$, $f_2(2 \mid 2) = \dfrac{f(2, 2)}{f_1(2)} = \dfrac{0.2}{0.6} = \dfrac{2}{6} = \dfrac{1}{3}$

Illustration 2 : During the throw of two dice simultaneously, if X assumes the value on the first die and Y assumes the value on the second die then find (i) P(X + Y = 6) (ii) P(X = Y) (iii) P(X + Y < 4)

Solution : We have

$X = 1, 2, 3, 4, 5, 6;\ Y = 1, 2, 3, 4, 5, 6$

$\therefore \quad f(x, y) = P(X = x, Y = y) = \frac{1}{36}$ for each x and y.

(i) $P(X + Y = 6) = f(1, 5) + f(2, 4) + f(3, 3) + f(4, 2) + f(5, 1)$

$$= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

$$= \frac{5}{36}$$

(ii) $P(X = Y) = f(1, 1) + f(2, 2) + f(3, 3) + f(4, 4) + f(5, 5) + f(6, 6)$

$$= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

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

(iii) $P(X + Y < 4) = f(1, 2) + f(2, 1)$

$$= \frac{1}{36} + \frac{1}{36} = \frac{2}{36} = \frac{1}{18}$$

Illustration 3 : The joint probability distribution is given in the following table. Find the value of *k*.

Y \ X	0	1	2
1	$2k$	k	$4k$
2	$3k$	$4k$	$2k$

Solution : We know that $\sum_{\forall x} \sum_{\forall y} f(x, y) = 1$

$\therefore \quad f(0, 1) + f(0, 2) + f(1, 1) + f(1, 2) + f(2, 1) + f(2, 2) = 1$

$\therefore \quad 2k + 3k + k + 4k + 4k + 2k = 1$

$$\Rightarrow 16k = 1 \Rightarrow k = \frac{1}{16}$$

Illustration 4 : A bag contains 4 white, 2 red and 3 black balls. Let X be the number of red balls and Y be the number of black balls. If two balls are drawn at random then find the joint probability distribution.

Solution : We have $X = 0, 1, 2;\ Y = 0, 1, 2, 3;$

$f(x, y) = P(X = x, Y = y)$

$\therefore \quad f(0, 0) = P(\text{all 2 balls are white})$

$$= \frac{{}^4C_2}{{}^9C_2} = \frac{1}{6}$$

$f(0, 1) = P(\text{1 black and 1 is white ball})$

$$= \frac{{}^3C_1 \cdot {}^4C_1}{{}^9C_2} = \frac{1}{3}$$

$f(0, 2) = P(\text{2 black balls}) = \frac{{}^3C_2}{{}^9C_2} = \frac{1}{12}$

$f(0, 3) = 0$ (Only two balls are drawn)

$f(1, 0) = P(\text{1 red ball and 1 white})$

$$= \frac{{}^2C_1 \cdot {}^4C_1}{{}^9C_2} = \frac{2}{9}$$

$f(1, 1) = P(\text{1 red and 1 black ball})$

$$= \frac{{}^2C_1 \cdot {}^3C_1}{{}^9C_2} = \frac{1}{6}$$

$f(1, 2) = 0$

$f(1, 3) = 0$

$f(2, 0) = P(\text{2 red balls}) = \frac{{}^2C_2}{{}^9C_2} = \frac{1}{36}$

$f(2, 1) = 0$

$f(2, 2) = 0$

$f(2, 3) = 0$

We summarize these as

Y \ X	0	1	2
0	$\frac{1}{6}$	$\frac{2}{9}$	$\frac{1}{36}$
1	$\frac{1}{3}$	$\frac{1}{6}$	0
2	$\frac{1}{12}$	0	0
3	0	0	0

Illustration 5 : Out of five cars, two have tyre problems and one has brake problem and two are in good running condition. Two cars are required for the journey. If two cars are selected among five at random and if X denotes the number with tyre problem, Y denotes with brake problem then find

(i) The joint probability distribution

(ii) The probability of 0 or 1 total problem among the selection.

(iii) The marginal probability function of X and Y.

(iv) The conditional probability distribution of X given Y = 0.

Solution :

(i) We have X = 0, 1, 2; Y = 0, 1

$\therefore$ $f(0, 0)$ = P(Both cars have no problem)

$$= \frac{^2C_2}{^5C_2} = \frac{1}{10}$$

$f(0, 1)$ = P(1 with brake problem, 1 with no problem) $= \frac{^1C_1 \cdot {^2C_1}}{^5C_2} = \frac{1}{5}$

Similarly

$f(1, 0) = \frac{^2C_1 \cdot {^2C_1}}{^5C_2} = \frac{2}{5}$ (2 are with tyre problem, 2 with no problem)

$$f(1, 1) = \frac{^2C_1 \cdot {^1C_1}}{^5C_2} = \frac{1}{5}$$

$f(2, 0) = \frac{^2C_2}{^5C_2} = \frac{1}{10}$; $f(2, 1) = 0$ (since only 2 cars are selected)

We summarize it as

Y \ X	0	1	2
0	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{1}{10}$
1	$\frac{1}{5}$	$\frac{1}{5}$	0

(ii) The probability of 0 or 1 total problem

$= f(0, 0) + f(0, 1) + f(1, 0)$

$$= \frac{1}{10} + \frac{1}{5} + \frac{2}{5} = \frac{7}{10}$$

(iii) The marginal probabilities $f_1(x)$ of X and $f_2(y)$ of Y are as given in the following table :

Y \ X	0	1	2	$f_2(y)$
0	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{1}{10}$	$\frac{6}{10}$
1	$\frac{1}{5}$	$\frac{1}{5}$	0	$\frac{2}{5}$
$f_1(x)$	$\frac{3}{10}$	$\frac{3}{5}$	$\frac{1}{10}$	Total 1

(iv) The conditional probability distribution of X given Y = 0.

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)}$$

For $x = 0$, $f_1(0 \mid 0) = \frac{f(0, 0)}{f_2(0)} = \frac{1/10}{6/10} = \frac{1}{6}$

For $x = 1$, $f_1(1 \mid 0) = \frac{f(1, 0)}{f_2(0)} = \frac{2/5}{6/10} = \frac{2}{3}$

For $x = 2$, $f_1(2 \mid 0) = \dfrac{f(2, 0)}{f_2(0)} = \dfrac{1/10}{6/10} = \dfrac{1}{6}$

Illustration 6 : The joint probability distribution for two random variables X and Y is given by $f(x, y) = \dfrac{x+y+1}{15}$, where $x = 0, 1, 2$ and $y = 0, 1$. Find the marginal probability distribution of X and Y. Also find conditional probability distribution of Y given X = 1.

Solution : We have $f(x, y) = \dfrac{x+y+1}{15}$, $x = 0, 1, 2$; $y = 0, 1$

$\therefore$ We summarize this in the table

Y \ X	0	1	2	$f_2(y)$
0	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{6}{15}$
1	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{9}{15}$
$f_1(x)$	$\frac{3}{15}$	$\frac{5}{15}$	$\frac{7}{15}$	Total 1

The marginal probability distributions $f_1(x)$ and $f_2(y)$ are given in the above table.

The conditional probability distribution of Y given X = 1 is

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)}$$

$$\therefore \quad f_1(0 \mid 1) = \frac{f(1, 0)}{f_1(1)} = \frac{2/15}{5/15} = \frac{2}{5}$$

$$f_2(1 \mid 1) = \frac{f(1, 1)}{f_1(1)} = \frac{3/15}{5/15} = \frac{3}{5}$$

Note : From the above table

$$f(0, 0) = \frac{1}{15} \neq f_1(0)\, f_2(0),$$

$$f(1, 1) = \frac{3}{15} \neq f_1(1)\, f_2(1) = \frac{5}{15} \cdot \frac{9}{15} = \frac{45}{225}$$

$\therefore$ The variables X and Y are dependent.

2.5.4 Continuous Random Variables :

During random experiment there exist continuous random variables like length, height, width measurements of some objects, volume of fluids, etc. Let X and Y be such continuous random variables then (X, Y) is a two dimensional random variable. Let $a \le X \le b$ and $c \le Y \le d$ then the **joint probability density** is denoted by $f(x, y)$ and the probability is given by

$$\int_a^b \int_c^d f(x, y)\, dy\, dx$$

provided that $f(x, y) \ge 0$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$.

For $X \le x$ and $Y \le y$ the **joint cumulative distribution function** is denoted by $F(x, y)$ and is given by

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

2.5.5 Marginal Probability Function for C.R.V. :

Let X and Y be two continuous random variables having joint probability density function is $f(x, y)$.

(a) The probability density function of x, that is the individual distribution function, denoted by $f_1(x)$ is called **marginal density function of X** and given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Likewise, the marginal density function of Y is defined and denoted by

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

The marginal distribution function is given by

$$F_1(x) = F(x, \infty) = \int_{-\infty}^{x} \left\{ \int_{-\infty}^{\infty} f(x, y)\, dy \right\} dx$$

$$F_2(y) = F(\infty, y) = \int_{-\infty}^{y} \left\{ \int_{-\infty}^{\infty} f(x, y)\, dx \right\} dy$$

2.5.6 Conditional Probability Density Function and Independent Variables :

Let X and Y be two continuous random variables. The **conditional probability density function of X given Y** is defined as

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)}, \text{ provided } f_2(y) \neq 0$$

Similarly, the **conditional probability density function of Y given X** is defined as

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)}, \text{ provided } f_1(x) \neq 0$$

The two random variables X and Y are **independent** if, $f(x, y) = f_1(x)\, f_2(y)$ or

$$f_1(x \mid y) = f_1(x) \text{ or } f_2(y \mid x) = f_1(x)$$

Illustration 1 : The joint probability density function of two random variables X and Y is given by

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

Find

(i) the value of k (ii) the marginal density function of X and Y (iii) the conditional probability density function.

Also check whether X and Y are independent.

Solution :

(i) We know that for joint probability density function

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

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

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

$$\Rightarrow k \int_0^1 [2x + 4]\, dx = 1$$

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

$$\Rightarrow k[1 + 4] = 1 \Rightarrow k = \frac{1}{5}$$

(ii) The marginal density function of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$= \int_0^2 \frac{x + 2y}{5}\, dy = \frac{1}{5}\left[xy + y^2 \right]_0^2$$

$$= \frac{1}{5}[2x + 4],\ 0 < x < 1$$

and $f_1(x) = 0$ elsewhere

The marginal density function of Y is given by

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

$$= \int_0^1 \frac{x + 2y}{5}\, dx = \frac{1}{5}\left[\frac{x^2}{2} + 2xy \right]_0^1$$

$$= \frac{1}{5}\left[\frac{1}{2} + 2y \right] = \frac{1 + 4y}{10},\ 0 < y < 2$$

and $f_2(y) = 0$ elsewhere

(iii) The conditional probability density function of X given Y is

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{(x + 2y)10}{10(1 + 4y)} = \frac{x + 2y}{1 + 4y};$$

$0 < x < 1;\ 0 < y < 2$

and $f_1(x \mid y) = 0$ for $x \leq 0$ or $x \geq 1$ and $0 < y < 2$

The conditional probability density function of Y given X is

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{(x + 2y)5}{10(2x + 4)} = \frac{x + 2y}{4(x + 2)}$$

for $0 < x < 1,\ 0 < y < 2$

and $f_2(y \mid x) = 0$ for $y \le 0$ or $y \ge 2$ and $0 < x < 1$

Now we have,

$$f_1(x) \cdot f_2(y) = \frac{2(x+2)}{5} \cdot \frac{(1+4y)}{10}$$

$$= \frac{(x+2)(1+4y)}{25} \ne f(x, y)$$

Thus X and Y are not independent.

Illustration 2 : Let the joint probability density function be given by

$$f(x, y) = \begin{cases} ke^{-x-2y} & ; x > 0, y > 0 \\ 0 & ; \text{elsewhere} \end{cases}$$

(i) Find k

(ii) Find the probability for $x < 2$ and $y > 2$

(iii) Find joint cumulative distribution function

(iv) Find marginal density function of X and Y

(v) Find conditional probability density of Y given X

Solution :

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

$$\therefore \int_0^{\infty}\int_0^{\infty} ke^{-x-2y}\, dx\, dy = 1$$

$$\Rightarrow k\int_0^{\infty}\left[-e^{-x-2y}\right]_0^{\infty} dy = 1 \Rightarrow k\int_0^{\infty} -e^{-2y}\, dy = 1$$

$$\Rightarrow k\left[\frac{e^{-2y}}{-2}\right]_0^{\infty} = 1 \Rightarrow k\left[0 + \frac{1}{2}\right] = 1 \Rightarrow k = 2$$

(ii) We have

$$\int_0^{2}\int_2^{\infty} 2e^{-x-2y}\, dy\, dx = \int_0^{2} 2\left[\frac{e^{-x-2y}}{-2}\right]_2^{\infty} dx$$

$$= \int_0^{2}\left[0 + e^{-x-4}\right] dx$$

$$= \left[-e^{-x-4}\right]_0^{2} = -e^{-6} + e^{-4}$$

(iii) $F(x, y) = \int_0^{x}\int_0^{y} 2e^{-x-2y}\, dy\, dx$ for $x > 0, y > 0$

$$= \int_0^{x} 2\left[\frac{e^{-x-2y}}{-2}\right]_0^{y} dx = \int_0^{x}\left[-e^{-x-2y} + e^{-x}\right] dx$$

$$= \left[e^{-x-2y} - e^{-x}\right]_0^{x} = e^{-x-2y} - e^{-x} - e^{-2y} + 1$$

$$= e^{-x}(e^{-2y} - 1) - (e^{-2y} - 1) = (e^{-x} - 1)(e^{-2y} - 1);$$

$x > 0, y > 0$ and $F(x, y) = 0$ elsewhere

(iv) Marginal density function of X on Y is

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$= \int_0^{\infty} 2e^{-x-2y}\, dy \text{ for } x > 0$$

$$= 2\left[\frac{e^{-x-2y}}{-2}\right]_0^{\infty} = 0 + e^{-x} = e^{-x} \text{ for } x > 0$$

and $f_1(x) = 0$ elsewhere

(v) Conditional probability density of Y on X is

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{2e^{-x-2y}}{e^{-x}} = 2e^{-2y}$$

for $x > 0, y > 0$

and $f_2(y \mid x) = 0$ for $y \le 0$ and $x > 0$.

EXERCISE

1. Three coins are tossed simultaneously. Obtain the probability distribution of the number of heads in a toss.

   Ans. : $\left(0, \frac{1}{8}\right), \left(1, \frac{3}{8}\right), \left(2, \frac{3}{8}\right), \left(3, \frac{1}{8}\right)$

2. A coin is weighted so that $P(H) = \frac{2}{3}$ and $P(T) = \frac{1}{3}$ is tossed three times. Obtain probability distribution of the number of heads.

   Ans. : $\left(0, \frac{1}{27}\right), \left(1, \frac{6}{27}\right), \left(2, \frac{12}{27}\right), \left(3, \frac{8}{27}\right)$

3. A random variable X is given by the following probability distribution.

X	–1	0	1	2	3
$f(x)$	0.2	0.1	k	$2k$	0.1

(i) Find the value of k (ii) Find $P(1 \le x \le 3)$ and $P(x > 1)$. **Ans. :** (i) 0.2, (ii) 0.7, 0.5.

4. A random variable X has the following distribution :

X	–2	–1	0	1	2
$f(x)$	0.2	0.1	0.1	0.3	0.3

Find distribution function.

Ans. : F(x) = 0.2, 0.3, 0.4, 0.7, 1

5. An ice cream seller can earn Rs. 200 per day if it is a hot day. He looses Rs. 60 per day, if it is a cool day. What is his expectation if the probability of hot day is 0.4 ? **Ans. :** Rs. 44

6. A die is thrown twice. Getting a number greater than 4 is taken as a success. Obtain probability distribution of the number of successes.

[Hint : $p = \frac{2}{6}$, $q = \frac{2}{3}$]

Ans. : $\left(0, \frac{4}{9}\right), \left(1, \frac{4}{9}\right), \left(2, \frac{1}{9}\right)$

7. For the pdf $f(x) = kx, \quad 0 < x < 1$
$= 0 \quad$, elsewhere

Find k. **Ans. :** $k = 2$

8. For the pdf $f(x) = 0.5\,(x + 1) \quad , -1 \le x \le 1$
$= 0 \quad$, elsewhere

Find distribution function.

Ans. : $F(x) = \frac{(x+1)^2}{4} \quad , -1 \le x \le 1$

$= 0 \quad , x < -1$

$= 1 \quad , x > 1$

9. Which of the following cannot be probability distribution ?

(i) $f(x) = \frac{x}{5}; x = 1, 2, 3, 4$

(ii) $f(x) = \frac{x^2}{6}, x = 4, 3, 2$

(iii) $f(x) = \frac{x}{10}, x = 10, 4, 15, 2$ **Ans. :** A

10. For discrete random variables X and Y, the joint probability function $f(x, y)$ is given in the following table :

Y \ X	0	1	2
0	0.1	0.1	0.2
1	0.2	0.3	0.1

Find

(i) $P(X + Y < 3)$

(ii) The marginal probability function of X and Y

(iii) The conditional probability of X given Y =

Ans. : (i) 0.4, (ii) $f_1(x) = 0.3, 0.4, 0.3$ for $x =$ 0, 1, 2 respectively and $f_2(y) = 0.4, 0.6$ for $y =$ 0, 1 respectively. (iii) $f_1(0 \mid 0) = 0.25$, $f_1(1 \mid 0) = 0.25$, $f_1(2 \mid 0) = 0.5$

11. The joint probability density function for two continuous random variables X and Y is given by

$$f(x, y) = \begin{cases} 6e^{-2x-3y} & , \text{for } x > 0, y > 0 \\ 0 & , \text{elsewhere} \end{cases}$$

(i) Find marginal probability of X and Y

(ii) Are X and Y independent ?

Ans. : (i) $f_1(x) = \begin{cases} 2e^{-2x} & , \text{for } x > 0 \\ 0 & , \text{elsewhere} \end{cases}$

$f_2(y) = \begin{cases} 3e^{-3y} & , \text{for } y > 0 \\ 0 & , \text{elsewhere} \end{cases}$

(ii) X and Y are independent.

* * *