8 Probability Distributions

Let's Study

- Random variables
- Types of random variables
- Probability distribution of a random variable
  - Discrete random variable
    - Probability mass function
    - Cumulative distribution function
    - Expected value and variance
  - Continuous random variable
    - Probability density function
    - Cumulative distribution function
  - Binomial distribution
  - Bernoulli trial
- Mean and variance of Binomial distribution
- Poisson distribution

Let's Recall

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

Let's Learn

8.1 Random variables:

We have already studied random experiments and sample spaces corresponding to random experiments. As an example, consider the experiment of tossing two fair coins. The sample space corresponding to this experiment contains four points, namely {HH, HT, TH, TT}.

We have already learnt to construct the sample space of any random experiment. However, the interest is not always in a random experiment and its sample space. We are often not interested in the outcomes of a random experiment, but only in some number obtained from the outcome. For example, in case of the experiment of tossing two fair coins, our interest may be only in the number of heads when two coins are tossed. In general, it is possible to associate a unique real number with very possible outcome of a random experiment. The number obtained from an outcome of a random experiment can take different values for different outcomes. This is why such a number is a variable. The value of this variable depends on the outcome of the random experiment, and is therefore called a random variable. A random variable is usually denoted by capital letters like X, Y, Z, ...

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

- When we throw two dice, there are 36 possible outcomes, but if we are interested in the sum of the numbers on the two dice, then there are only 11 different possible values, from 2 to 12.
- If we toss a coin 10 times, then there are $2^{10} = 1024$ possible outcomes, but if we are interested in the number of heads among the 10 tosses of the coin, then there are only 11 different possible values, from 0 to 10.
- In the experiment of randomly selecting four items from a lot of 20 items that contains 6 defective items, the interest is in the number of defective items among the selected four items. In this case, there are only 5 different possible outcomes, from 0 to 4.

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

A random variable is formally defined as follows:

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

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

We often use the abbreviation *r.v.* for random variable.

Consider an experiment where three seeds are sown in order to find how many of them germinate. Every seed will either germinate or will not germinate. Let us use the letter Y when a seed germinates. The sample space of this experiment can then be written as $S = \{YYY, YYN, YNY, NYY, YNN, NYN, NNY, NNN\}$ and $n(S) = 8$.

None of these outcomes is a number. We shall try to represent every outcome by a number. Consider the number of times the letter Y appears is a possible outcome and denote it by X. Then, we have $X(YYY) = 3$, $X(YYN) = X(YNY) = X(NYY) = 2$, $X(YNN) = X(NYN) = X(NNY) = 1$, $X(NNN) = 0$.

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

A random variable is denoted by a capital letter, like X and Y. A particular value taken by the random variable is denoted by the small letter x. Note that x is a real number and the set of all possible outcomes corresponding to a particular value x of X is denoted by the event $[X = x]$. For example, in the experiment of three seeds, the random variable X has four possible values, namely 0, 1, 2, 3. The four events are then defined as follows.

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

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

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

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

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

A sample space need not be finite. Consider, for example, the experiment of tossing a coin until a head is obtained. The sample space for this experiment is $S = \{H, TH, TTH, TTTH, ...\}$. Note that S contains an unending sequence of tosses required to get a head. Here, S is countably infinite. The random variable.

$X : S \rightarrow$ R, denoting the number of tosses required to get a head, has the range $\{1, 2, 3, ...\}$ which is also countably infinite.

8.2 Types of Random Variables:

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

8.2.1 Discrete Random Variable:

Definition : A random variable is a discrete random variable if its possible values form a countable set, which may be finite or infinite.

The values of a discrete random variable are usually denoted by non-negative integers, that is, 0, 1, 2, ... Examples of discrete random variables include the number of children in a family, the number of patients in a hospital ward, the number of cars sold by a dealer, and so on.

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

8.2.2 Continuous Random Variable

Definition: A random variable is a continuous random variable if its possible values form an interval of real numbers.

A continuous random variable has uncountably infinite possible values and these values form an interval of real numbers. Examples of continuous random variables include heights of trees in a forest, weights of students in a class, daily temperature of a city, speed of a vehicle, and so on.

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

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

8.3 Probability Distribution of a Discrete Random Variable

Let us consider the experiment of throwing two dice and noting the numbers on the uppermost faces of the two dice. The sample space of this experiment is $S = \{(1,1), (1,2), ..., (6,6)\}$ and $n(S) = 36$.

Let X denote the sum of the two numbers in a single throw. Then the set of possible values of X is $\{2, 3, ..., 12\}$. Further,

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

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

. . .

. . .

. . .

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

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

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

This leads to the following results.

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

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

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

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

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

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

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

Definition : The probability distribution of a discrete random variable X is defined by the following system. Let the possible values of X be denoted by x_1, x_2, x_3, ..., and the corresponding probabilities be denoted by p_1, p_2, p_3, ... Then, the set of ordered pairs $\{(x_1, p_1), (x_2, p_2), (x_3, p_3)$, ...} is called the **probability distribution of the random variable X**.

For example, consider the coin-tossing experiment where the random variable X is defined as the number of tosses required to get a head. Let the probability of getting head be 't' and that of not getting head be $1 - t$. The possible values of X are given by the set of natural numbers, {1, 2, 3, ...} and $P[X = i] = (1 - t)^{i-1}t$, for $i = 1,2,3,....$. This result can be verified by noting that if head is obtained for the first time on the ith toss, then the first $i - 1$ tosses have resulted in tail. In other words, $[X = i]$ represents the event of having $i - 1$ tails followed by the first head on the ith toss.

$p_i = P[X = x_i]$ for $i = 1, 2, 3, \ldots\ldots.$.

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

The probability distribution of a discrete random variable is sometimes presented in a tabular form as follows.

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

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

8.3.1 Probability Mass Function (p. m.f.)

The probability p_i of X taking the value x_i *is* sometimes a function of x_i for all possible values of X. In such cases, it is sufficient to specify all possible values of X and the function that gives probablilities of these values. Such a function is called the probability mass function (p. m. f.) of the discrete random variable X.

For example, consider the coin-tossing experiment where the random variable X is defined as the number of tosses required to get a head. Let probability of getting a head be 't' and that of not getting a head be $1 - t$. The possible values of X are given by the set of natural numbers {1,2,3,.......} and $P[X = i] = (1 - t)^{i-1}t$, for $i = 1,2,3,\ldots\ldots$ is a function of i.

The probability mass function (p. m. f.) of a discrete random variable is defined as follows.

Definition. Let the possible values of a discrete random variable X be denoted by $x_1, x_2, \ldots\ldots$, with the corresponding probabilities $p_i = P[X = x_i]$, $i = 1, 2, \ldots$ If there is a function f such that $f(x_i) = p_i = P[X = x_i]$ for all possible values of X, then f is called the probability mass function (p. m. f.) of X.

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

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

Note that $P[X = x] = \binom{4}{x}\left(\frac{1}{2}\right)^4$, for $x = 0$, 1, 2, 3, 4, where $\binom{4}{x}$ is the number of ways of getting x heads in 4 tosses.

8.3.2 Cumulative Distribution Function (c. d. f.)

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

The cumulative distribution function (c. d. f.) of a discrete random variable is defined as follows.

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

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

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

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

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

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

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

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

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

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

Let X denote the number of heads obtained in two tosses, Then X is a discrete random variable and its values for the possible outcomes of the experiment are obtained as follows.

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

Let Y denote the number of heads minus the number of tails in two tosses. Then Y is also a discrete random variable and its values for the possible outcomes of the experiment are obtained as follow.

$$Y(HH) = 2, Y(HT) = Y(TH) = 0,$$

$$Y(TT) = -2.$$

Let $Z = \frac{\text{number of heads}}{\text{number of tails} + 1}$. Then Z is also a discrete random variable and its values for the possible outcomes of the experiment are obtained as follows.

$$Z(HH) = 2, Z(HT) = Z(TH) = \frac{1}{2}, Z(TT) = 0.$$

These examples show that it is possible to define many discrete random variables on the same sample space. Possible values of a discrete random variable can be positive or negative, integer or fraction, and so on, as long as they are countable.

SOLVED EXAMPLES

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

Solution: X is a number whose value depends on the outcome of a random experiment. Therefore, X is a random variable. Since the sample space of the experiment has only 8 possible outcomes, X is a discrete random variable. Now, the sample space of the experiment is S = $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$. The values of X corresponding to theses outcomes of the experiment are as follows.

$$\begin{aligned} X(HHH) &= 2 \times 3 \\ &= 6 \\ X(HHT) &= X(HTH) \\ &= X(THH) \\ &= 2 \times 2 - 1.50 \times 1 \\ &= 2.50 \\ X(HTT) &= X(THT) \\ &= X(TTH) \\ &= 2 \times 1 - 1.50 \times 2 \\ &= -1.00 \end{aligned}$$

$$X(TTT) = -1.50 \times 3 = -4.50$$

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

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

Solution: Let the balls in the bag be denoted by r, g_1, g_2. The sample space of the experiment is then given by $S = \{rr, rg_1, rg_2, g_1r, g_1g_1, g_1g_2, g_2r, g_2g_1, g_2g_2\}$.

Since X is the number of red balls, we have

$$X(\{rr\}) = 2$$

$$X(\{rg_1\}) = X(\{rg_2\}) = X(\{g_1r\}) = X(\{g_2r\}) = 1$$

$$X(\{g_1g_1\}) = X(\{g_1g_2\}) = X(\{g_2g_1\}) = X(\{g_2g_2\}) = 0$$

Thus X is a discrete random variable with possible values, 0, 1 and 2. The probability distribution of X is obtained as follows.

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

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

Solution: Let X denote the number of aces among the two cards drawn with replacement. Clearly, 0, 1 and 2 are the possible values of X. Since the draws are with replacement, the outcomes of the two draws are independent of each other. Also, since there are 4 aces in the deck of 52 cards, $P[\text{an ace}] = \frac{4}{52} = \frac{1}{13}$, and $P[\text{a non-ace}] = \frac{12}{13}$.

Then

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

$$P[X = 1] = P[\text{ace and non-ace}] + P[\text{non-ace and ace}] = \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} = \frac{24}{169}$$

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

The requaried probability distribution is then as follows.

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

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

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

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

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

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

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

The probability distribution of X is as shown in the following table.

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

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

Solution : Let X denote the number of doubles. The possible doubles in a single throw of a pair of dice are given by (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6).

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

$P[X = 0] = P$ [no doublet]

$= qqq = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{216}$

$P[X = 1] = P$ [one doublets]

$pqq + qpq + qqp$

$= 3pq^2 = \frac{75}{216}$

$P[X = 2] = P$ [two doublets]

$= ppq + pqp + qpp$

$= 3p^2q = \frac{15}{216}$

$P[X = 3] = P$ [three doublets]

$= ppp = \frac{1}{216}$

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

x	0	1	2	3	4
$P[X = x]$	0.1	k	$2k$	$2k$	k

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

Solution: The table gives a probability distribution and therefore.

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

That is, $0.1 + k + 2k + 2k + k = 1$

That is, $6k = 0.9$ Therefore $k = 0.15$.

(i) $k = 0.15$

(ii) $P[X < 2] = P[X = 0] + P[X = 1] = 0.1 + k = 0.1 + 0.15 = 0.25$

(iii) $P[X \geq 3] = P[X = 3] + P[X = 4] = 2k + k = 3(0.15) = 0.45$

(iv) $P[1 \leq X < 4] = P[X = 1] + P[X = 2] + P[x = 3] = k + 2k + 2k = 5k = 5(0.15) = 0.75$.

v) $F(2) = P[X \leq 2] = P[X = 0] + P[x = 1] + P[X = 2] = 0.1 + k + 2k = 0.1 + 3k = 0.1 + 3(0.15) = 0.1 + 0.45 = 0.55$.

8.3.3 Expected value and variance of a random variable.

In many problems, the interest is in some feature of a random variable computed from its probability distribution. Some such numbers are mean, variance and standard deviation. We shall discuss mean and variance in this section. Mean is a measure of location in the sense that it is the average value of the random variable.

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

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

The mean or expected value of a random variable X is the sum of products of possible values of X and their respective probabilities.

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

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

The non-negative square root of Var(X) is called the standard deviation of the random variable X. That is, $\sigma = \sqrt{\text{Var}(X)}$.

Another formula to find the variance of a random variable. We can also use the simplified form of Var $(X) = \sum_{i=1}^{n} x^2{}_i p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$

or Var(X) = $E(X^2) - [E(X)]^2$.

SOLVED EXAMPLES

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

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

x_i	p_i	$x_i \; p_i$	$x^2{}_i p_i$
0	$\frac{1}{8}$	0	0
1	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{8}$
2	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{12}{8}$
3	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{9}{8}$

Then $E(X) = \sum_{i=1}^{n} x_i \, p_i = \frac{12}{8} = 1.5.$

$$\text{Var}(X) = \sum_{i=1}^{n} x^2{}_i \, p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \frac{24}{8} - (1.5)^2$$

= 3 – 2.25 = 0.75.

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

Solution : The sample space of the experiment consists of 36 elementary events in the form of ordered pairs (x_i, y_i), where x_i = 1, 2, 3, 4, 5, 6 and y_i = 1, 2, 3, 4, 5, 6. The random variable X has the possible values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.

x_i	P_i	$x_i p_i$	$x^2{}_i p_i$
2	1/36	2/36	4/36
3	2/36	6/36	18/36
4	3/36	12/36	48/36
5	4/36	20/36	100/36
6	5/36	30/36	180/36
7	6/36	42/36	294/36
8	5/36	40/36	320/36
9	4/36	36/36	324/36
10	3/36	30/36	300/36
11	2/36	22/36	242/36
12	1/36	12/36	144/36
		$\sum_{i=1}^{n} x_i \, p_i = \frac{252}{36} = 7$	$\sum_{i=1}^{n} x^2{}_i \, p_i = \frac{1974}{36} = 54.83$

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

$$\text{Var}(X) = \sum_{i=1}^{n} x^2{}_i \, p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = 5.83$$

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

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

Let X denote the selected number. Then X is a random variable with possible values 1, 2, 3, ..., 15. Each of these numbers is equiprobable.

Therefore, $P(1) = P(2) = P(3) = ... = P(15) = \frac{1}{15}$.

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

$$+ 15 \times \frac{1}{15} = (1 + 2 + ... + 15) \times \frac{1}{15}$$

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

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

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

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

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

$$= 82.67 - 64 = 18.67.$$

Ex. 4. Two cards are drawn without replacement from a well shuffled pack of 52 cards. Find the mean and variance of the number of kings drawn.

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

$$P\text{ (both cards are not king)} = \frac{\binom{48}{2}}{\binom{52}{2}} = \frac{48 \times 47}{52 \times 51} = \frac{188}{221}.$$

P (one card is king and other card is not king)

$$= \frac{\binom{4}{1} \times \binom{48}{1}}{\binom{52}{2}} = \frac{4 \times 48 \times 2}{52 \times 51} = \frac{32}{221}$$

$$P\text{ (both cards are king)} = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{4 \times 3}{52 \times 51} = \frac{1}{221}.$$

Therefore,

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

$$= 0 \times \frac{188}{221} + 1 \times \frac{32}{221} + 2 \times \frac{1}{221} = \frac{34}{221},$$

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

$$= \left(0^2 \times \frac{188}{221} + 1^2 \times \frac{32}{221} + 2^2 \times \frac{1}{221}\right) - \left(\frac{34}{221}\right)^2$$

$$= \frac{36}{221} - \frac{1156}{48841} = \frac{6800}{48841} = 0.1392$$

EXERCISE 8.1

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

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

3. Determine whether each of the following is a probability distribution. Give reasons for your answer.

(i)

x	0	1	2
$P(x)$	0.4	0.4	0.2

(ii)

x	0	1	2	3	4
$P(x)$	0.1	0.5	0.2	–0.1	0.3

(iii)

x	0	1	2
$P(x)$	0.1	0.6	0.3

(iv)

z	3	2	1	0	–1
$P(z)$	0.3	0.2	0.4	0.05	0.05

(v)

y	-1	0	1
$P(y)$	0.6	0.1	0.2

(vi)

x	0	1	2
$P(x)$	0.3	0.4	0.2

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

5. Find the probability distribution of the number of successes in two tosses of a die if success is defind as getting a number greater than 4.

6. A sample of 4 bulbs is drawn at random with replacement from a lot of 30 bulbs which includes 6 defectives bulbs. Find the probability distribution of the number of defective bulbs.

7. A coin is blased so that the head is 3 times as likely to occur as tail. Find the probability distribution of number of tails in two tosses.

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

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

Determine (i) k, (ii) $P(X < 3)$,

(iii) $P(0 < X < 3)$, (iv) $P(X > 4)$.

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

x	-2	-1	0	1	2
$P(x)$	0.2	0.3	0.1	0.15	0.25

10. Find expected value and variance of X, the number on the uppermost face of a fair die.

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

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

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

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

15. A class has 15 students whose ages are 14, 17, 15, 14, 21, 17, 19, 20, 16, 18, 20, 17, 16, 19 and 20 years. If X denotes the age of a randomly selected student, find the probability distribution of X. Find the mean and variance of X.

16. 70% of the members favour and 30% oppose a proposal in a meeting. The random variable X takes the value 0 if a member opposes the proposal and the value 1 if a member is in favour. Find $E(X)$ and Var(X).

8.4 Probability Distribution of a Continuous Random Variable

The possible values of a continuous random variable form an interval of real numbers. The probability distribution of a continuous random variable is represented by a continuous function called the probability density function (p. d. f.). A continuous random variable also has a cumulative distribution function (c. d. f.).

Suppose the possible values of the continuous random variable X form the interval (a, b), where a and b are real numbers such that a < b. The interval (a, b) is called the support of the continuous random variable X.

We shall now define the probability density function (p. d. f.) and cumulative distribution fucntions (c. d. f.) of a continuous random variable.

8.4.1 Probability Density Function (p. d. f.)

Let X be a continuous random variable with the interval (a, b) as its support. The probability density function (p. d. f.) of X is an integrable function f that satisfies the following conditions.

1. $f(x) \geq 0$ for all $x \in (a, b)$.

2. $\int_a^b f(x)\,dx = 1$

3. For any real numbers c and d such that

$$a \leq c < d \leq b,$$

$$P[X \in (c, d)] = \int_c^d f(x)\,dx$$

It is easy to notice that the p. d. f. of a continuous random variable is different from the p. m. f. of a discrete random variable. Both the p. m. f. and p. d. f. are positive at possible values of the random variable. However, the p. d. f. is positive over an entire interval that is, over an uncountably infinite set of points.

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

The cumulative distribution function (p. d. f.) of a continuous random variable is defined as follows.

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

$$F(x) = 0 \text{ for all } x \leq a,$$

$$= \int_a^x f(x)\,dx \text{ for all } x \geq a.$$

Note: The c. d. f. of a continuous random variable is a non-decreasing continuous function. The c. d. f. of a discrete random variable is a step function, while the c. d. f. of a continuous random variable is a continuous function.

SOLVED EXAMPLES

Ex. 1. Let X be a continuous random variable with probability function (p. d. f.) $f(x) = 3x^2$ for $0 < x < 1$. Can we claim that $f(x) = P[X = x]$?

Solution. Note, for example, that $f(0.9) = 3(0.9)^2 = 2.43$, which is not a probability. This example shows that the p. d. f. of a continuous random variable does not represent the probability of a possible value of the random variable. In case of a continuous random variable, the probability of an interval is obtained by integrating the p. d. f. over the specified interval. In this case, the probability is given by the area under the curve of the p. d. f. over the interval.

Let us now verify whether f is a valid probability density function (p. d. f). This is done through the following steps.

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

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

Therefore, the function $f(x) = 3x^2$ for $0 < x < 1$ is a proper probability density function. Also, for real numbers c and d such that $0 \leq c < d \leq 1$, note that $P[c < X < d] = \int_c^d f(x)\,dx = \int_c^d 3x^2 dx = \left[x^3\right]_c^d = d^3 - c^3 > 0$.

What is the probability that X falls between 1/2 and 1?

That is, what is $P(1/2 < X < 1]$?

Take $c = 1/2$ and $d = 1$ in the above integral to obtain $P[1/2 < X < 1] = 1 - (1/2)^3 = 1 - 1/8 = 7/8$.

What is $P(X = 1/2)$?

It is easy to see that the probability is 0. This is so because $\int_c^d f(x)\,dx = \int_{1/2}^{1/2} 3x^2 dx = (1/2)^3 - (1/2)^3 = 0$.

As a matter of fact, the probability that a continuous random variable X takes any specific value x is 0. That is, when X is a continuous random variable, $P[X = x] = 0$ for all real x.

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

Solution. Not that the internal of the p. d. f. over the support of the random variable must be 1. That is $\int_0^c f(x)\,dx = 1$. That is, $\int_0^c \frac{x^3}{4} dx = 1$. But, $\int_0^c \frac{x^3}{4} dx = \frac{c^4}{16}$. Since this integral must be equal to I, we have $\frac{c^4}{16} = 1$, or equivalently $c^4 = 16$, so that $c = 2$ since c must be positive.

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

$f(x) = 3x^2$ for $0 < x < 1$.

Find the cumulative distribution function $F(x)$.

Solution : $F(x) = \int_0^x 3x^2 dx = \left[x^3 - (0)^3\right]_0^x = x^3$ for $x \in (0,1)$.

Therefore,

$$f(x) = \begin{cases} 0 & \text{for } x \le 0, \\ x^3 & \text{for } 0 < x < 1, \\ 1 & \text{for } x \ge 1. \end{cases}$$

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

$f(x) = \frac{x^3}{4}$ for $0 < x < 2$.

Find the cumulative distribiuton function of X.

Solution : $F(x) = \int_0^x f(x)\,dx = \int_0^x \frac{x^3}{4} dx = \frac{1}{4}\left[\frac{x^4}{4}\right]_0^x = \frac{1}{16}\left[x^4 - 0\right]_0^x = \frac{x^4}{16}$ for $0 < x < 2$.

Therefore,

$$F(x) = \begin{cases} 0 & \text{for } x < 0, \\ \frac{x^4}{16} & \text{for } 0 \le x < 2, \\ 1 & \text{for } x \ge 2. \end{cases}$$

Ex. 5. Verify whether each of the following functions is p. d. f. of a continuous r. v. X.

(i) $f(x) = \begin{cases} e^{-x} & \text{for } 0 < x < \infty \\ 0 & \text{otherwise.} \end{cases}$

(ii) $f(x) = \begin{cases} \frac{x}{2} & \text{for } -2 < x < 2, \\ 0 & \text{otherwise.} \end{cases}$

Solution :

(i) $e^{-x} \ge 0$ for all real values of x, because $e > 0$. Therefore, $e^{-x} > 0$ for $0 < x < \infty$.

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

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

(ii) $f(x)$ is negative for $-2 < x < 0$, and therefore $f(x)$ is not a valid p. d. f.

Ex. 6. Find k if the following function is the p. d. f. of a r. v. X.

$$f(x) = \begin{cases} kx^2(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

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

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

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

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

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

$$\therefore k \times 12 = 1$$

$$\therefore k = 12$$

Ex. 7. In each of the following cases, find (a) $P(X < 1)$ and (b) $P(|X| < 1)$.

(i) $f(x) = \begin{cases} \frac{x^2}{18} & \text{for } -3 < x < 3, \\ 0 & \text{otherwise} \end{cases}$

(ii) $f(x) = \begin{cases} \frac{x+2}{18} & \text{for } -2 < x < 4, \\ 0 & \text{otherwise} \end{cases}$

Solution:

(i) (a) $P(X < 1) = \int_{-3}^{1} \frac{x^2}{18} dx = \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-3}^{1}$

$= \frac{1}{54}[1 - (-3)^3] = \frac{1}{54}(1 + 27) = \frac{28}{54}$

$= \frac{14}{27}$

(b) $P(|X| < 1) = P(-1 < X < 1) = \int_{-1}^{1} \frac{x^2}{18} dx$

$= \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-3}^{1} = \frac{1}{54}[1 - (-1)^3]$

$= \frac{1}{54}(1 + 1) = \frac{2}{54} = \frac{1}{27}$

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

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

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

(b) $P(|X| < 1) = P(-1 < X < 1)$

$= \int_{-1}^{1} \frac{x+2}{18} dx = \frac{1}{18}\left[\frac{x^2}{2} + 2x\right]_{-1}^{1}$

$= \frac{1}{18}\left\{\left(\frac{1}{2}+2\right) - \left(\frac{1}{2} - 2\right)\right\}$

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

Ex. 8. Find the c. d. f. $F(x)$ associated with the p. d. f. $f(x)$ or r. v. X where

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

Solution: Since $f(x)$ is a p. d. f., the c. d. f. is given by

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

$= \left[3x - 2x^3\right]_0^x = 3x - 2x^3$ for $0 < x < 1$.

Therefore,

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

EXERCISE 8.2

1. Check whether each of the following is a p. d. f.

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

   ii) $f(x) = 2$ for $0 < x < 1$.

2. The following is the p. d. f. of a r. v. X.

   $f(x) = \begin{cases} \frac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise} \end{cases}$

   Find (i) $P(x < 1.5)$, (ii) $P(1 < x < 2)$, (iii) $P(x > 2)$.

3. It is felt that error in measurement of reaction temperature (in celesus) in an experiment is a continuous r. v. with p. d. f.

   $f(x) = \begin{cases} \frac{x^3}{64} & \text{for } 0 \le x \le 4, \\ 0 & \text{otherwise} \end{cases}$

   (i) Verify whether $f(x)$ is a p. d. f.

   (ii) Find $P(0 < x \le 1)$.

   (iii) Find probability that X is between 1 and 3.

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

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

   Also find $P\left[\frac{1}{4} < X < \frac{1}{2}\right]$

(ii) $f(x) = \begin{cases} kx(1-x) & \text{for } 0 < x < 1, \\ 0 & \text{otherwise} \end{cases}$

Also find (a) $P\left[\frac{1}{4} < X < \frac{1}{2}\right]$,

(b) $P\left[X < \frac{1}{2}\right]$

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

$$f(x) = \begin{cases} 0.5x & \text{for } 0 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

Calculate (i) $P(X \le 1)$, (ii) $P(0.5 \le X \le 1.5)$, (iii) $P(X \ge 1.5)$.

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

$$f(x) = \begin{cases} \frac{1}{5} & \text{for } 0 \le x \le 5, \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that (i) waiting time is between 1 and 3 minutes, (ii) waiting time is more than 4 minutes.

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

$$f(x) = \begin{cases} k(4-x^2) & \text{for } -2 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

Compute (i) $P(X > 0)$, (ii) $P(-1 < x < 1)$, (iii) $P(X < -0.5 \text{ or } X > 0.5)$.

8. Following is the p. d. f. of a continuous r. v. X.

$$f(x) = \begin{cases} \frac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise} \end{cases}$$

(i) Find expression for the c. d. f. of X. (ii) Find $F(x)$ at $x = 0.5, 1.7$, and 5.

9. The p. d. f. of a continuous r. v. X is

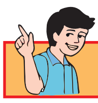
$$f(x) = \begin{cases} \frac{3x^2}{8} & \text{for } 0 < x < 2, \\ 0 & \text{otherwise} \end{cases}$$

Determine the c. d. f. of X and hence find (i) $P(X < 1)$, (ii) $P(X < -2)$, (iii) $P(X > 0)$, (iv) $P(1 < X < 2)$.

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

$$f(x) = \begin{cases} \frac{c}{x} & \text{for } 1 < x < 3, \\ 0 & \text{otherwise} \end{cases}, \; c > 0..$$

Find c, $E(X)$, and Var (X). Also find $F(x)$.



Let's Remember

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

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

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

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

Discrete random variable :

Let the possible values of a discrete random variable X be denoted by $x_1, x_2, \ldots$, with the corresponding probabilities $p_i = P[X = x_i]$, $i = 1, 2, \ldots$. If there is a function f such that $p_i = P[X = x]$, $f(x_i)$ for all possible values of X, then f is called the probability mass function (p. m. f.) of X.

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

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

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

$$\begin{aligned} F(x) &= P[X \le x] \\ &= \sum_{x_i \le x} P[X = x_i] \\ &= \sum_{x_i \le x} f(x_i). \end{aligned}$$

Expected Value or Mean of a Discrete r. v.

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

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

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

Variance of a Discrete r. v.

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

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

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

Another formula to find the variance of a random variable. We can also use the simplified form $\text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$, or

$$\text{Var}(X) = E(X^2) - [E(X)]^2,$$

where $E(X^2) = \sum_{i=1}^{n} x_i^2 p_i$.

Probability Density Function (p. d. f.)

Let X be a continuous random variable defined on the interval $S = (a, b)$. A non-negative integrable function f is called the probability density function (p. d. f.) of X if it satisfies the following conditions. (i) $f(x) > 0$ for all $x \in S$. (ii) The area under the curve $f(x)$ over S is 1. That is, $\int_S f(x)\, dx = 1$. (iii) The probability that X takes a value in some interval A is given by the integral of $f(x)$ over A. That is $P[X \in A] = \int_A f(x)\, dx$.

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

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

8.5 Binomial Distribution

Let's Recall

Many experiments are dichotomous in nature. An experiment is dichotomous if it has only two possible outcomes. For example, a tossed coin shows 'head' or 'tail', the result of a student is 'pass' or 'fail', a manufactured item is 'defective' or 'non-defective', the response to a question is 'yes' or 'no', an egg has 'hatched' or 'not hatched', the decision is 'yes' or 'no' etc. In such cases, it is customary to call one of the outcomes 'success' and the other 'failure'. For example, in tossing a coin, if the occurrence of head is considered success, then occurrence of tail is failure.

8.5.1 Bernoulli trial

An experiment that can result in one of two possible outcomes is called a dichotomous experiment. One of the two outcomes is called success and the other outcome is called failure.

Definition of Bernoulli Trial : A dichotomous experiment is called a Bernoulli trial.

Every time we toss a coin or perform a dichotomous experiment, we call it a trial. If a coin is tossed 4 times, the number of trials is 4, each having exactly two possible outcomes, namely success and failure. The outcome of any trial is independent of the outcome of other trials. In all such trials, the probability of success (and hence of failure) remains the same.

Definition of Sequence of Bernoulli Trials: A sequence of dichotomous experiments is called a sequence of Bernoulli trials if it satisfies the following conditions.

- The trials are independent.
- The probability of success remains the same in all trials.

The probability of success in a Bernoulli trial is denoted by p and the probability of failure is denoted by $q = 1 - p$. For example, if we throw a die and define success as getting an even number and failure as getting an odd number, we have a Bernoulli trial. Successive throws of the die are independent trials and form a sequence of Bernoulli trials. If the die is fair, then $p = 1/2$ and $q = 1 - p = 1 - 1/2 = 1/2$.

Ex. 1. Six balls are drawn successively from an urn containing 7 red and 9 black balls. Decide whether the trials of drawing balls are Bernoulli trials if, after each draw, the ball drawn is (i) replaced (ii) not replaced in the urn.

Solution.

(i) When the drawing is done with replacement, the probability of success (say, red ball) is $p = 7/16$ which remains the same for all six trials (draws). Hence, drawing balls with replacement are Bernoulli trials.

(ii) When the drawing is done without replacement, the probability of success (i.e. red ball) in first trial is 7/16. In second trial, it is 6/15 if first ball drawn is red and is 7/15 if first ball drawn in black, and so on. Clearly probability of success is not same for all trials, and hence the trials are not Bernoulli trials.

8.5.2 Binomial distribution

Consider a sequence of Bernoulli trials. where we are interested in the number successes, X, in a given number of trials, n. The total number of possible outcomes in 2^n. We want to find the number of outcomes that result in X successes in order to find the probability of getting X successes in n Bernoulli trials. We use the binomial theorem and derive a formula for counting the number of favorable outcomes for all possible values of X.

Consider the situation where we are interested in one success in 6 trials. Clearly, six possible cases are {*SFFFFF*, *FSFFFF*, *FFSFFF*, *FFFSFF*, *FFFFSF*, *FFFFFS*}. Counting the number of outcomes with two successes is possible by listing favorable outcomes. However, it will be time-consuming to list all of these outcomes. As a result, calculating probabilities of getting 0, 1, 2, ..., 6 successes in 6 trials will be time consuming. Lengthy calculations that require listing of all favorable outcomes for calculating probabilities of number of successes in n Bernoulli trials can be avoided by deriving a formula.

For illustrating this formula, let us take the experiment made up of three Bernoulli trials with probablities p and $q = 1 - p$ of success and failure, respectively, in each trial. The sample space of the experiment is the set.

$$S = \{SSS, SSF, SFS, SFF, FSF, FFS, FFF\}.$$

The number of successes is a random variable X and can take values 0, 1, 2, and 3. The probability distribution of the number of successes is as follows.

$$P(X = 0) = P \text{ (no success)}$$
$$= P(\{FFF\}) = P(F)\,P(F)\,P(F)$$
$$= q \,.\, q \,.\, q = q^3,$$
since trials are independent.

$$P(X = 1) = P\text{(one success)}$$
$$= P(\{SFF, FSF, FFS\})$$
$$= P(\{SFF\}) + P(\{FSF\}) + P(\{FFS\})$$
$$= P(S)P(F)P(F) + P(F)P(S)P(F) + P(F)P(F)P(S)$$
$$= p.q.q + q.p.q + q.q.p = 3pq^2$$

$P(X = 2)$ = P(two successes)

= $P(\{SSF, SFS, FSS\})$

= $P(\{SSF\}) + P(\{SFS\}) + P(\{FSS\})$

= $P(S)P(S)P(F) + P(S)P(F)P(S) + P(F)P(S)P(S)$

= $p.p.q + p.q.p + q.p.p = 3p^2q$

and $P(X = 3)$ = P(three successes)

= $P(\{SSS\})$

= $P(S).P(S).P(S) = p^3$.

Thus, the probability distribution of X is as shown in the following table.

x	0	1	2	3
$P(x)$	q^3	$3q^2p$	$3qp^2$	p^3

Note that the binomial expansion of $(q+p)^3$ is $q^3+3p^2q+3qp^2+p^3$ and the probabilities of 0, 1, 2 and 3 successes are respectively the 1st, 2nd, 3rd and 4th term in the expansion of $(q + p)^3$. Also, since $q + p = 1$, it follows that the sum of these probabilities, as expected, is 1. Thus, we may conclude that in an experiment of n Bernoulli trials, the probabilities of 0, 1, 2, ..., n successes can be obtained as 1st, 2nd, ..., $(n + 1)$th terms in the expansion of $(q + p)^n$. To prove this assertion (result), let us find the probability of x successes in an experiment of n Bernoulli trials.

In an experiment with n trials, if there are x successes (S), there will be $(n - x)$ failure (F). Now x successes (S) and $(n - x)$ failure (F) can be obtained in $\dfrac{n!}{x!(n-x)!}$ ways. In each of these ways, the probabilitiy of x succcesses and $(n - x)$ failure, that is, P(x successes). P($n - x$ failure) is given by

$(P(S).P(S)$ ... $P(S)$ x times) $(P(F).P(F)$ ... $P(F)$ $(n - x)$ times)

= (p.p.p...p x times) (q.q.q...q (n − x) times)

= $p^x q^{n-x}$.

Thus, the probability of x successes in n Bernoulli trials is

P(x successes out of n trials)

$$\frac{n!}{x!(n-x)!}\,p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

Note that

P(x successes) = $\binom{n}{x} p^x q^{n-x}$ is the $(x + 1)$th term in the binomial expansion of $(q + p)^n$.

Thus, the probability distribution of number of successes in an experiment consisting of n Bernoulli trials can be obtained by the binomial expansion of $(q + p)^n$. This distribution of X successes in n Bernoulli trials can be written as follows.

X	0	1	2	...
P(X)	$\binom{n}{0} p^0 q^n$	$\binom{n}{1} p^1 q^{n-1}$	$\binom{n}{2} p^2 q^{n-2}$	...
X	x	...	n	
P(X)	$\binom{n}{x} p^x q^{n-x}$	...	$\binom{n}{n} p^n q^0$	

The above probability distribution is known as the **binomial distribution** with parameters n and p. We can find the complete probability distribution of the random variable using the given values of n and p. The fact that the r. v. X follows the binomial distribution with parameters n and p is written in short as $X \sim B(n, p)$ and read as "X follows the binomail distribution with parameters n and p."

The probability of x successes $P(X = x)$ is also denoted by $P(x)$ and is given by $P(x) = {}^nC_x\, q^{n-x} p^x, x = 0, 1, ...n; (q = 1 - p)$.

This $P(x)$ is the probability mass function of the binomial distribution.

A binomial distribution with n Bernoulli trials and p as probability of success in each trial is denoted by B (n,p) or written as $X \sim B(n,p)$.

Let us solve some examples.

Ex. 2. If a fair coin is tossed 10 times, find the probability of obtaining.

(i) exactly six heads

(ii) at least six heads

(iii) at most six heads

Solution: The repeated tosses of a coin are Bernoulli trails. Let X denote the number of heads in an experiment of 10 trials.

Clearly, $X \sim B(n,p)$ with $n = 10$ and $p = \frac{1}{2}$ and $q = 1 - p = \frac{1}{2}$.

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = x) = \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}$$

(i) Exactly six successes means $X = 6$.

$$P(X = 6) = \binom{10}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^{10-6}$$

$$= \frac{10!}{6!(10-6)!}\left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^4$$

$$= \frac{10\times9\times8\times7}{4\times3\times2\times1}\left(\frac{1}{2}\right)^{10}$$

$$= \frac{105}{512}$$

(ii) At least six successes means $X \geq 6$.

$$P(X \geq 6) = P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$$

$$= \binom{10}{6}\left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^4 + \binom{10}{7}\left(\frac{1}{2}\right)^7 \times \left(\frac{1}{2}\right)^3 + \binom{10}{8}\left(\frac{1}{2}\right)^8 \times \left(\frac{1}{2}\right)^2 + \binom{10}{9}\left(\frac{1}{2}\right)^9 \times \left(\frac{1}{2}\right)^1 + \binom{10}{10}\left(\frac{1}{2}\right)^{10} \times \left(\frac{1}{2}\right)^0$$

$$= \frac{10\times9\times8\times7}{4\times3\times2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10\times9\times8}{3\times2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10\times9}{2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10}$$

$$= (210 + 120 + 45 + 10 + 1) \times \frac{1}{1024}$$

$$= \frac{386}{1024} = \frac{193}{512}$$

(iii) At most six successes means $X \leq 6$.

$$P(X \leq 6) = 1 - P(X > 6)$$

$$= 1 - [P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)]$$

$$= 1 - [\binom{10}{7}\left(\frac{1}{2}\right)^7 \times \left(\frac{1}{2}\right)^3 + \binom{10}{8}\left(\frac{1}{2}\right)^8 \times \left(\frac{1}{2}\right)^2 + \binom{10}{9}\left(\frac{1}{2}\right)^9 \times \left(\frac{1}{2}\right)^1 + \binom{10}{10}\left(\frac{1}{2}\right)^{10} \times \left(\frac{1}{2}\right)^0 +]$$

$$= 1 - \left[\frac{10\times9\times8}{3\times2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10\times9}{2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10}\right]$$

$$= 1 - \left[(120 + 45 + 10 + 1) \times \frac{1}{1024}\right]$$

$$= 1 - \frac{176}{1024} = 1 - \frac{11}{64} = \frac{53}{64}$$

Ex. 3. The eggs are drawn successively with replacement from a lot containing 10% defective eggs. Find the probability that there is at least one defective egg.

Solution: Let X denote the number of defective eggs in the 10 eggs drawn. Since the drawing is done with replacement, the trials are Bernoulli trials.

Probability of success $= \frac{1}{10}$

$\therefore \quad P = \frac{1}{10}, q = 1 - p = 1 - \frac{1}{10} = \frac{9}{10}$, and

$n = 10$.

$X \sim B\left(10, \frac{1}{10}\right)$.

$$P(X = x) = \binom{10}{x}\left(\frac{1}{10}\right)^{x}\left(\frac{9}{10}\right)^{10-x}$$

Here, $X \geq 1$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X = 0) \\ &= 1 - \binom{10}{0}\left(\frac{1}{10}\right)^{0}\left(\frac{9}{10}\right)^{10} \\ &= 1 - 1 \times \left(\frac{9}{10}\right)^{10} \\ &= 1 - \left(\frac{9}{10}\right)^{10} \end{aligned}$$

8.5.3 Mean and Variance of Binomial Distribution

Let $X \sim B(n,p)$. The mean or expected value of X is denoted by μ. It is also called expected value of X and is denoted by $E(X)$ and given by $\mu = E(X) = np$. The variance is denoted by $\text{Var}(X)$ and is given by $\sigma^2 = \text{Var}(X) = npq$.

Ex. 4. If $X \sim B(10, 0.4)$, then find $E(X)$ and $\text{Var}(X)$.

Solution: Here $n = 10$, $p = 0.4$, $q = 1 - p = 1 - 0, 4 = 0.6$.

$E(X) = np = 10(0.4) = 4$

$\text{Var}(X) = npq = 10(0.4)(0.6) = 2.4$

Ex. 5. Let the p.m. f. of the r. v. X be $P(X = x) = \binom{4}{x}\left(\frac{5}{9}\right)^{x}\left(\frac{4}{9}\right)^{4-x}$, for $x = 0, 1, 2, 3, 4$. Find $E(X)$ and $\text{Var}(X)$.

Solution: X follows the binomial distribution with $n = 4$, $p = 5/9$ and $q = 4/9$.

$$\text{Var}(X) = npq = 4\left(\frac{5}{9}\right)\left(\frac{4}{9}\right) = 80/81.$$

Ex. 6. If $E(X) = 6$ and $\text{Var}(X) = 4.2$, find n and p.

Solution: $E(x) = 6$. Therefore, $np = 6$. $\text{Var}(X) = 4.2$.

Therefore, $npq = 4.2$.

$$\frac{npq}{np} = \frac{4.2}{6}$$

$\therefore \quad q = 0.7, p = 1 - q = 1 - 0.7 = 0.3$

$np = 6 \ \therefore \ n \times 0.3 = 6 \ \therefore \ n = \frac{6}{0.3} = 20$

EXERCISE 8.3

1. A die is thrown 4 times. If 'getting an odd number' is a success, find the probability of

   (i) 2 successes (ii) at least 3 successes (iii) at most 2 successes.

2. A pair of dice is thrown 3 times. If getting a doublet is considered a success, find the probability of two successes.

3. There are 10% defective items in a large bulk of items. What is the probability that a sample of 4 items will include not more than one defective item?

4. Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. Find the probability that

   (i) all the five cards are spades.

   (ii) only 3 cards are spades.

   (iii) none is a spade.

5. The probability that a bulb produced by a factory will fuse after 200 days of use is 0.2. Let X denote the number of bulbs (out of 5) that fuse after 200 days of use. Find the probability of

   (i) $X = 0$, (ii) $X \leq 1$, (iii) $X > 1$, (iv) $X \geq 1$.

6. 10 balls are marked with digits 0 to 9. If four balls are selected with replacement. What is the probability that none is marked 0?

7. In a multiple choice test with three possible answers for each of the five questions, what is the probability of a candidate getting four or more correct answers by random choice?

8. Find the probability of throwing at most 2 sixes in 6 throws of a single die.

9. Given that $X \sim B(n,p)$,

   (i) if $n = 10$ and $p = 0.4$, find $E(X)$ and $\text{Var}(X)$.

   (ii) if $p = 0.6$ and $E(X) = 6$, find n and $\text{Var}(X)$.

   (iii) if $n = 25$, $E(X) = 10$, find p and Var (X).

   (iv) if $n = 10$, $E(X) = 8$, find $\text{Var}(X)$.

8.6 Poisson Distribution

Poisson distribution is a discrete probability distribution that gives the probability of number of occurrences of an event in a fixed interval of time, if these events occur with a known average rate and are independent of the time since the last occurrence. For instance, suppose someone receives 4 emails per day on an average. There will be, however, variation in the number of emails, sometimes more, sometimes fewer, once in a while no email at all. The Poisson distribution was first introduced by Simeon Denis Poisson (1781-1840) and published in 1837. The work focused on certain random variables N that count the number of discrete occurrences of an event that take place during a time-interval of given length.

Definition: A discrete random variable X is said to have the Poisson distribution with parameter $m > 0$, if its p.m. is given by

$$P(X = x) = \frac{e^{-m}m^{x}}{x!} = 0, 1, 2, ...$$

Note.

(i) we use the notation $X \sim P(m)$ to show that X follows the Poisson distribution with parameter m.

(ii) Observe that $P(x) > 0$ for all non-negative integers x and $\sum_{x=0}^{\infty} P(x) = 1$.

(iii) For the Poisson distribution Mean = $E(X) = m$ and Variance = $\text{Var}(X) = m$.

(iv) When n is very large and p is very small in the binomial distribution, then X follows the Poisson distribution with parameter m = np.

SOLVED EXAMPLES

Ex. 1. If $X \sim P(m)$ with $m = 5$ and $e^{-5} = 0.0067$, then find

(i) $P(X = 5)$, (ii) $P(X \geq 2)$.

Solution: $P(X = x) = \dfrac{e^{-m}m^{x}}{x!}$, for $x = 0, 1, 2, ...$.

(i) Here $m = 5$ and $x = 5$

$$P(X = 5) = \frac{e^{-5}5^{5}}{5!} = \frac{0.0067 \times 3125}{5 \times 4 \times 3 \times 2 \times 1} = 0.1745$$

(ii)
$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \left[\frac{e^{-5}5^{0}}{0!} + \frac{e^{-5}5^{1}}{1!}\right] \\ &= 1 - \left[\frac{0.0067 \times 1}{1} + \frac{0.0067 \times 5}{1}\right] \\ &= 1 - [0.0067 + 0.0335] \\ &= 1 - 0.0402 \\ &= 0.9598 \end{aligned}$$

Ex. 2. If $X \sim P(m)$ with $P(X = 1) = P(X = 2)$, then find the mean and $P(X = 2)$ given that $e^{-2} = 0.1353$.

Solution: Since $P(X = 1) = P(X = 2)$,

$$\therefore \quad \frac{e^{-m}m^1}{1!} = \frac{e^{-m}m^2}{2!}$$

$$\therefore \quad m = 2$$

Then

$$P(X = 2) = \frac{e^{-2}2^2}{2!} = \frac{0.1353 \times 4}{2} = 0.2706.$$

Ex. 3. In a town, 10 accidents takes place in the span of 50 days. Assuming that the number of accidents follows Poisson distribution, find the probability that there will be 3 or more accidents on a day.

(given that $e^{-0.2} = 0.8187$)

Solution: Here $m = \dfrac{10}{50} = 0.2$, and hence $X \sim P(m)$ with $m = 0.2$. The p. m. f. of X is $P(X = x) = \dfrac{e^{-m}m^x}{x!}$, $x = 0, 1, 2, ...$

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= \quad 1 - \left[\frac{e^{-0.2}(0.2)^0}{0!} + \frac{e^{-0.2}(0.2)^1}{1!} + \frac{e^{-0.2}(0.2)^2}{2!}\right]$$

$$= 1 - \left[\frac{0.8187 \times 1}{1} + \frac{0.8187 \times 0.2}{1} + \frac{0.8187 \times 0.004}{2}\right]$$

$$= 1 - [0.8187 + 0.16374 + 0.016374]$$

$$= 1 - 0.9988 = 0.0012$$

EXERCISE 8.4

1. If X has Poisson distribution with m = 1, then find $P(X \leq 1)$ given $e^{-1} = 0.3678$.

2. If $X \sim P(1/2)$, then find $P(X = 3)$ given $e^{-0.5} = 0.6065$.

3. If X has Poisson distribution with parameter m and $P(X = 2) = P(X = 3)$, then find $P(X \geq 2)$. Use $e^{-3} = 0.0497$.

4. The number of complaints which a bank manager receives per day follows a Poisson distribution with parameter m = 4. find the probability that the manager receives (i) only two complaints on a given day, (ii) at most two complaints on a given day. Use $e^{-4} = 0.0183$.

5. A car firm has 2 cars, which are hired out day by day. The number of cards hired on a day follows Poisson distribution with mean 1.5. Find the probability that (i) no car is used on a given day, (ii) some demand is refused on a given day, given $e^{-1.5} = 0.2231$.

6. Defects on plywood sheet occur at random with the average of one defect per 50 sq. ft. Find the probability that such a sheet has (i) no defect, (ii) at least one defect. Use $e^{-1} = 0.3678$.

7. It is known that, in a certain area of a large city, the average number of rats per bungalow is five. Assuming that the number of rats follows Poisson distribution, find the probability that a randomly selected bungalow has (i) exactly 5 rats (ii) more than 5 rats (iii) between 5 and 7 rats, inclusive. Given $e^{-5} = 0067$.

Let's Remember

Trials of a random experiment are called Bernoulli trials, if they satisfy the following conditions.

- Each trail has two possible outcomes, success and failure.
- The probability of success remains the same in all trials.

Thus, the probability of getting x successes in n Bernoulli trials is

$$P(x \text{ successes in } n \text{ trials}) = \binom{n}{x} p^x \times q^{n-x}$$

$$= \frac{n!}{x! \times (n-x)!} p^x \times q^{n-x}.$$

Clearly, P (x successess) = $\binom{n}{x} p^x \times q^{n-x}$ is the $(x + 1)^{\text{th}}$ term in the binomial expansion of $(q + p)^n$.

Let $X \sim B(n,p)$. Then the mean or expected value of r. v. X is denoted by μ. It is also denoted by $E(X)$ and is given by $\mu = E(X) = np$. The variance is denoted by $\text{Var}(X)$ and is given by $\text{Var}(X) = npq$.

A discrete random variable X is said to follow the Poisson distribution with parameter $m > 0$ if its p. m. f. is given by

$$P(X = x) = \frac{e^{-m} m^x}{x!}, \; x = 0, 1, 2, ...$$

Note.

1. We use the notation $X \sim P(m)$. That is, X follows Poisson distribution with parameter m.
2. Observe that $P(x) > 0$ for all non-negative integers x and $\sum_{x=0}^{\infty} \frac{e^{-m} m^x}{x!} = 1$
3. For the Poisson distribution, Mean $E(X) = m$ and Variance $= \text{Var}(X) = m$.
4. If n is very large and p is very small then X follows Poisson distribution with $m = np$.

MISCELLANEOUS EXERCISE - 8

I) Choose the correct alternative.

1. F(x) is c.d.f. of discrete r.v. X whose p.m.f. is given by $P(x) = k\binom{4}{x}$, for $x = 0, 1, 2, 3, 4$ & $P(x) = 0$ otherwise then F(5) = ...

   (a) $\frac{1}{16}$ (b) $\frac{1}{8}$ (c) $\frac{1}{4}$ (d) 1

2. F(x) is c.d.f. of discrete r.v. X whose distribution is

X_i	–2	–1	0	1	2
P_i	0.2	0.3	0.15	0.25	0.1

   then F(–3) = ...

   (a) 0 (b) 1 (c) 0.2 (d) 0.15

3. X : is number obtained on upper most face when a fair die is thrown then E(X) = ...

   (a) 3.0 (b) 3.5 (c) 4.0 (d) 4.5

4. If p.m.f. of r. v. X is given below.

x	0	1	2
$P(x)$	q^2	$2pq$	p^2

   then $\text{Var}(x)$ = ...

   (a) p^2 (b) q^2 (c) pq (d) $2pq$

5. The expected value of the sum of two numbers obtained when two fair dice are rolled is ....

   (a) 5 (b) 6 (c) 7 (d) 8

6. Given p.d.f. of a continuous r.v. X as

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

   $= 0$ otherwise then

   $F(1) =$

   (a) $\frac{1}{9}$ (b) $\frac{2}{9}$ (c) $\frac{3}{9}$ (d) $\frac{4}{9}$

7. X is r.v. with p.d.f. $f(x) = \frac{k}{\sqrt{x}}$, $0 < x < 4$

   = 0 therwise

   then E (x) = ...

   (a) $\frac{1}{3}$ (b) $\frac{4}{3}$ (c) $\frac{2}{3}$ (d) 1

8. If $X \sim B(20, \frac{1}{10})$ then E (x) = ...

   (a) 2 (b) 5 (c) 4 (d) 3

9. If E (x) = m and Var (x) = m then X follows ....

   (a) Binomial distribution
   (b) Possion distribution
   (c) Normal distribution
   (d) none of the above

10. If E (x) > Var (x) then X follows ....
   (a) Binomial distribution
   (b) Possion distribution
   (c) Normal distribution
   (d) none of the above

II) Fill in the blanks.

1. The values of discrete r.v. are generally obtained by ....

2. The values of continuous r.v. are generally obtained by ....

3. If X is discrete random variable takes the values $x_1, x_2, x_3, \ldots x_n$ then $\sum_{i=1}^{n} P(x_i) =$ ....

4. If F(x) is distribution function of discrete r.v.x with p.m.f. P(x) = $\frac{x-1}{3}$ for x = 1, 2, 3, & P(x) = 0 otherwise then F(4) = ...

5. If F(x) is distribution function of discrete r.v. X with p.m.f. P(x) = $k\binom{4}{x}$ for x = 0, 1, 2, 3, 4, and P(x) = 0 otherwise then F(−1) = ...

6. E(x) is considered to be .... of the probability distribution of x.

7. If x is continuous r.v. and F(x_i) = P(X ≤ x_i) = $\int_{-\infty}^{x_i} f(x)\,dx$ then F(x) is called ....

8. In Binomial distribution probability of success .... from trial to trial.

9. In Binomail distribution if n is very large and probability success of p is very small such that $np = m$ (constant) then .... distribution is appplied.

III) State whether each of the following is True or False.

1. If P(X = x) = $k\binom{4}{x}$ for x = 0,1,2,3,4, then F(5) = $\frac{1}{4}$ when F(x) is c.d.f.

2.

X	−2	−1	0	1	2
P(X = x)	0.2	0.3	0.15	0.25	0.1

If F(x) is c.d.f. of discrete r.v. X then F(−3) = 0.

3. X is the number obtained on upper most face when a die is thrown then E(x) = 3.5.

4. If p.m.f. of discrete r.v. X is

X	0	1	2
$P(X = x)$	q^2	$2pq$	p^2

then E(x) = 2p.

5. The p.m.f. of a r.v. X is

P(x) = $\frac{2x}{n(n+1)}$, x = 1, 2, ... n

= 0 otherwise,

Then

E(x) = $\frac{2n+1}{3}$

6. If $f(x) = k\,x\,(1-x)$ for $0 < x < 1$
   = 0 otherwise then
   k = 12

7. If X ~ B(n,p) and n = 6 & P(x = 4) = P(x = 2) then p = $\frac{1}{2}$.

8. If r.v. X assumes values 1,2,3, ... n with equal probabilities then E(x) = $\frac{(n+1)}{2}$

9. If r.v. X assumes the values 1,2,3, ..., 9 with equal probabilities, E(x) = 5.

IV) Solve the following problems.

PART - I

1. Identify the random variable as discrete or continuous in each of the following. Identify its range if it is discrete.

   (i) An economist is interested in knowing the number of unemployed graduates in the town with a population of 1 lakh.

   (ii) Amount of syrup prescribed by a physician.

   (iii) A person on high protein diet is interested in the weight gained in a week.

   (iv) Twelve of 20 white rats available for an experiment are male. A scientist randomly selects 5 rats and counts the number of female rats among them.

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

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

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

   (i) Determine the value of k.

   (ii) Find $P(X \leq 4)$, $P(2 < X < 4)$, $P(X \geq 3)$.

3. Following is the probability distribution of a r. v. X.

x	−3	−2	−1	0	1	2	3
$P(X = x)$	0.05	0.1	0.15	0.20	0.25	0.15	0.1

   Find the probability that

   (i) X is positive.

   (ii) X is non-negative

   (iii) X is odd.

   (iv) X is even.

4. The p. m. f of a r. v. X is given by

$$P(X = x) = \begin{cases} \binom{5}{x}\dfrac{1}{2^5}, & x = 0,1,2,3,4,5. \\ 0 & \text{otherwise} \end{cases}$$

   Show that $P(X \leq 2) = P(X \geq 3)$.

5. In the following probability distribution of a r. v. X.

x	1	2	3	4	5
$P(x)$	1/20	3/20	a	$2a$	1/20

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

6. A fair coin is tossed 4 times. Let X denote the number of heads obtained. Identify the probability distribution of X and state the formula for p. m. f. of X.

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

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

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

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

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

x	−3	−2	−1	0	1	2	3	4
$F(x)$	0.1	0.3	0.5	0.65	0.75	0.85	0.9	1

   Find the probability distribution of X and $P(-1 \leq X \leq 2)$.

10. Find the expected value and variance of the r. v. X if its probability distribution is as follows.

   (i)

x	1	2	3
$P(X = x)$	1/5	2/5	2/5

(ii)

x	-1	0	1
$P(X=x)$	1/5	2/5	2/5

(iii)

x	1	2	3	...	n
$P(X=x)$	$1/n$	$1/n$	$1/n$	...	$1/n$

(iv)

x	0	1	2	3	4	5
$P(X=x)$	1/32	5/32	10/32	10/32	5/32	1/32

11. A player tosses two coins. He wins Rs. 10 if 2 heads appear, Rs. 5 if 1 head appears, and Rs. 2 if no head appears. Find the expected value and variance of winning amount.

12. Let the p. m. f. of the r. v. X be

$$P(x) = \begin{cases} \dfrac{3-x}{10} & \text{for } x = -1, 0, 1, 2. \\ 0 & \text{otherwise} \end{cases}$$

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

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

$$f(x) = \begin{cases} k(4-x^2) & \text{for } -2 \le x \le 2. \\ 0 & \text{otherwise} \end{cases}$$

Compute (i) $P(X > 0)$, (ii) $P(-1 < x < 1)$, (iii) $P(X < 0.5 \text{ or } X > 0.5)$

14. The p. d. f. of the r. v. X is given by

$$f(x) = \begin{cases} \dfrac{1}{2a} & \text{for } 0 < x < 2a. \\ 0 & \text{otherwise} \end{cases}$$

Show that $P(X < a/2) = P(X > 3a/2)$.

15. Determine k if the p. d. f. of the r. v. is

$$f(x) = \begin{cases} ke^{-\theta x} & \text{for } 0 \le x < \infty. \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X > \frac{1}{\theta})$ and detemine M if $P(0 < X < M) = \frac{1}{2}$

16. The p. d. f. of the r. v. X is given by

$$f(x) = \begin{cases} \dfrac{k}{\sqrt{x}} & \text{for } 0 < x < 4. \\ 0 & \text{otherwise} \end{cases}$$

Determine k, the c. d. f. of X, and hence find $P(X \le 2)$ and $P(x \ge 1)$.

17. Let X denote the reaction temperature in celcius of a certain chemical process. Let X have the p. d. f.

$$f(x) = \begin{cases} \dfrac{1}{10} & \text{for } -5 \le x < 5. \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(X < 0)$.

PART - II

1. Let X ~ B(10,0.2). Find (i) P(X = 1) (ii) P(X ≥ 1) (iii) P(X ≤ 8).

2. Let X~B(n,p) (i) If n = 10 and E(X)= 5, find p and Var(X). (ii) If E(X)= 5 and Var(X) = 2.5, find n and p.

3. If a fair coin is tossed 4 times, find the probability that it shows (i) 3 heads, (ii) head in the first 2 tosses and tail in last 2 tosses.

4. The probability that a bomb will hit the target is 0.8. Find the probability that, out of 5 bombs, exactly 2 will miss the target.

5. The probability that a lamp in the classroom will burn is 0.3. 3 lamps are fitted in the classroom. The classroom is unusable if the number of lamps burning in it is less than 2. Find the probability that the classroom can not be used on a random occasion.

6. A large chain retailer purchases an electric device from the manufacturer. The manufacturer indicates that the defective rate of the device is 10%. The inspector of the retailer randomly selects 4 items from a shipment. Find the probability that the inspector finds at most one defective item in the 4 selected items.

7. The probability that a component will survive a check test is 0.6. Find the probability that exactly 2 of the next 4 components tested survive.

8. An examination consists of 5 multiple choice questions, in each of which the candidate has to decide which one of 4 suggested answers is correct. A completely unprepared student guesses each answer completely randomly. Find the probability that this student gets 4 or more correct answers.

9. The probability that a machine will produce all bolts in a production run within the specification is 0.9. A sample of 3 machines is taken at random. Calculate the probability that all machines will produce all bolts in a production run within the specification.

10. A computer installation has 3 terminals. The probability that anyone terminal requires attention during a week is 0.1, independent of other terminals. Find the probabilities that (i) 0 (ii) 1 terminal requires attention during a week.

11. In a large school, 80% of the students like mathematics. A visitor asks each of 4 students, selected at random, whether they like mathematics. (i) Calculate the probabilites of obtaining an answer yes from all of the selected students. (ii) Find the probability that the visitor obtains the answer yes from at least 3 students.

12. It is observed that it rains on 10 days out of 30 days. Find the probability that (i) it rains on exactly 3 days of a week. (ii) it rains on at most 2 days of a week.

13. If X follows Poisson distribution such that $P(X = 1) = 0.4$ and $P(X = 2) = 0.2$, find variance of X.

14. If X follows Poisson distribution with parameter m such that $\dfrac{P(X=x+1)}{P(X=x)} = \dfrac{2}{x+1}$ find mean and variance of X.

✦✦✦