



PROBABILITY

61.1 PROBABILITY

Probability is a concept which numerically measure the degree of uncertainty and therefore, of certainty of the occurrence of events.

If an event A can happen in m ways, and fail in n ways, all these ways being equally likely to occur, then the probability of the happening of A is

$$= \frac{\text{Number of favourable cases}}{\text{Total number of mutually exclusive and equally likely cases}} = \frac{m}{m+n}$$

and that of its failing is defined as $\frac{n}{m+n}$

If the probability of the happening = p
and the probability of not happening = q

then
$$p+q = \frac{m}{m+n} + \frac{n}{m+n} = \frac{m+n}{m+n} = 1 \text{ or } p+q = 1$$

For instance, on tossing a coin, the probability of getting a head is $\frac{1}{2}$.

61.2 DEFINITIONS

1. **Die :** It is a small cube. Dots are :: :: marked on its faces. Plural of the die is dice. On throwing a die, the outcome is the number of dots on its upper face.
2. **Cards :** A pack of cards consists of four suits *i.e.* Spades, Hearts, Diamonds and Clubs. Each suit consists of 13 cards, nine cards numbered 2, 3, 4, ..., 10, and Ace, a King, a Queen and a Jack or Knave. Colour of Spades and Clubs is black and that of Hearts and Diamonds is red. Kings, Queens, and Jacks are known as *face* cards.
3. **Exhaustive Events or Sample Space :** The set of all possible outcomes of a single performance of an experiment is exhaustive events or sample space. Each outcome is called a sample point. In case of tossing a coin once, $S = (H, T)$ is the *sample space*. Two outcomes Head and Tail constitute an exhaustive event because no other outcome is possible.
4. **Random Experiment :** There are experiments, in which results may be altogether different, even though they are performed under identical conditions. They are known as random experiments. Tossing a coin or throwing a die is random experiment.
5. **Trail and Event :** Performing a random experiment is called a trial and outcome is termed as event. Tossing of a coin is a trial and the turning up of head or tail is an event.
6. **Equally likely events:** Two events are said to be '*equally likely*', if one of them cannot be expected in preference to the other. For instance, if we draw a card from well-shuffled pack, we may get any card, then the 52 different cases are equally likely.
7. **Independent event :** Two events may be *independent*, when the actual happening of one does not influence in any way the probability of the happening of the other.

Example. The event of getting head on first coin and the event of getting tail on the second

- coin in a simultaneous throw of two coins are independent.
8. **Mutually Exclusive events:** Two events are known as *mutually exclusive*, when the occurrence of one of them excludes the occurrence of the other. For example, on tossing of a coin, either we get head or tail, but not both.
 9. **Compound Event :** When two or more events occur in composition with each other, the simultaneous occurrence is called a compound event. When a die is thrown, getting a 5 or 6 is a compound event.
 10. **Favorable Events :** The events, which ensure the required happening, are said to be favourable events. For example, in throwing a die, to have the even numbers, 2, 4 and 6 are favourable cases.
 11. **Conditional Probability :** The probability of happening an event A , such that event B has already happened, is called the conditional probability of happening of A on the condition that B has already happened. It is usually denoted by $P(A/B)$.
 12. **Odds in favour of an event and odds against an event**

If number of favourable ways = m , number of not favourable events = n

$$(i) \text{ Odds in favour of the event} = \frac{m}{n}, \text{ Odds against the event} = \frac{n}{m}.$$

13. **Classical Definition of Probability.** If there are N equally likely, mutually, exclusive and exhaustive of events of an experiment and m of these are favourable, then the probability of

the happening of the event is defined as $\frac{m}{N}$.

14. **Expected value.** If $p_1, p_2, p_3, \dots, p_n$ of the probabilities of the events $x_1, x_2, x_3 \dots x_n$ respectively the expected value

$$E(x) = p_1 x_1 + p_2 x_2 + p_3 x_3 + \dots + p_n x_n = \sum_{r=1}^n p_r x_r$$

Example 1. Find the probability of throwing

(a) 5, (b) an even number with an ordinary six faced die.

Solution. (a) There are 6 possible ways in which the die can fall and there is only one way of throwing 5.

$$\text{Probability} = \frac{\text{Number of favourable ways}}{\text{Total number of equally likely ways}} = \frac{1}{6} \quad \text{Ans.}$$

(b) Total number of ways of throwing a die = 6

Number of ways falling 2, 4, 6 = 3

$$\text{The required probability} = \frac{3}{6} = \frac{1}{2} \quad \text{Ans.}$$

Example 2. Find the probability of throwing 9 with two dice.

Solution. Total number of possible ways of throwing two dice

$$= 6 \times 6 = 36$$

Number of ways getting 9. i.e., (3 + 6), (4 + 5), (5 + 4), (6 + 3) = 4.

$$\therefore \text{The required probability} = \frac{4}{36} = \frac{1}{9}. \quad \text{Ans.}$$

Example 3. From a pack of 52 cards, one is drawn at random. Find the probability of getting a king.

Solution. A king can be chosen in 4 ways.

But a card can be drawn in 52 ways.

$$\therefore \text{the required probability} = \frac{4}{52} = \frac{1}{13} \quad \text{Ans.}$$

EXERCISE 61.1

1. In a class of 12 students, 5 are boys and the rest are girls. Find the probability that a student selected will be a girl. Ans. $\frac{7}{12}$
2. A bag contains 7 red and 8 black balls. Find the probability of drawing a red ball. Ans. $\frac{7}{15}$
3. Three of the six vertices of a regular hexagon are chosen at random. Find the probability that the triangle with three vertices is equilateral. Ans. $\frac{1}{10}$
4. What is the probability that a leap year, selected at random, will contain 53 Sundays. *(A.M.I.E., Dec. 2009, Summer 2001)* Ans. $\frac{2}{7}$

Fill in the blanks with appropriate correct answer

5. Chance of throwing 6 at least once in four throws with single dice is

$$(A.M.I.E., Summer 2000) \quad \text{Ans. } \frac{671}{1296}$$

6. A pair of fair dice is thrown and one die shows a four. The probability that the other die shows 5 is *(A.M.I.E., Summer 2000)* Ans. $\frac{1}{36}$

61.3 ADDITION LAW OF PROBABILITY

If A and B are two events associated with an experiment; then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Proof. Let m_1 , m_2 , and m be the number of favourable outcomes to the events A, B and $A \cap B$ respectively. The mutually exclusive outcomes in the sample space of the experiment be n.

$$P(A) = \frac{m_1}{n}, \quad P(B) = \frac{m_2}{n}, \quad P(A \cap B) = \frac{m}{n}$$

The favourable outcomes to the event A only = $m_1 - m$

The favourable outcomes to the event B only = $m_2 - m$

The favourable outcomes to the event $A \cap B$ only = m.

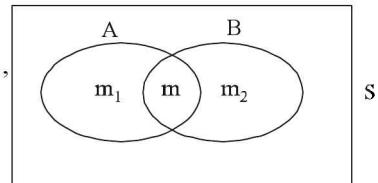
The favourable outcomes to the events A or B or both i.e.,

$$\begin{aligned} A \cup B &= m_1 - m + (m_2 - m) + m \\ &= m_1 + m_2 - m \end{aligned}$$

$$\text{So, } P(A \cup B) = \frac{m_1 + m_2 - m}{n}$$

$$= \frac{m_1}{n} + \frac{m_2}{n} - \frac{m}{n}$$

$$= P(A) + P(B) - P(A \cap B)$$



Theorem. If A and B are any two events prove that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and hence prove that if A, B and C are any three events.

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

(AMIETE, June 2010)

Note: Mutually Exclusive Events

Consider the case where two events A and B are not mutually exclusive. The probability of the event that either A or B or both occur is given as

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Example 4. An urn contains 10 black and 10 white balls. Find the probability of drawing two balls of the same colour.

Solution. Probability of drawing two black balls = $\frac{^{10}C_2}{^{20}C_2}$

\therefore Probability of drawing two white balls = $\frac{^{10}C_2}{^{20}C_2}$

\therefore Probability of drawing two balls of the same colour

$$= \frac{^{10}C_2}{^{20}C_2} + \frac{^{10}C_2}{^{20}C_2} = 2 \cdot \frac{^{10}C_2}{^{20}C_2} = 2 \cdot \frac{\frac{10 \times 9}{2 \times 1}}{\frac{20 \times 19}{2 \times 1}} = \frac{9}{19}$$

Ans.

Example 5. A bag contains four white and two black balls and a second bag contains three of each colour. A bag is selected at random, and a ball is then drawn at random from the bag chosen. What is the probability that the ball drawn is white?

Solution. There are two mutually exclusive cases,

(i) when the first bag is chosen, (ii) when the second bag is chosen.

Now the chance of choosing the first bag is $\frac{1}{2}$ and if this bag is chosen, the probability of drawing a white ball is $4/6$. Hence the probability of drawing a white ball from first bag is

$$\frac{1}{2} \times \frac{4}{6} = \frac{1}{3}$$

Similarly the probability of drawing a white ball from second bag is $\frac{1}{2} \times \frac{3}{6} = \frac{1}{4}$

Since the events are mutually exclusive the required probability = $\frac{1}{3} + \frac{1}{4} = \frac{7}{12}$ **Ans.**

61.4 MULTIPLICATION LAW OF PROBABILITY

If there are two independent events the respective probabilities of which are known, then the probability that both will happen is the product of the probabilities of their happening respectively.

$$P(AB) = P(A) \times P(B)$$

Proof. Suppose A and B are two independent events. Let A happen in m_1 ways and fail in n_1 ways.

$$\therefore P(A) = \frac{m_1}{m_1 + n_1}$$

Also let B happen in m_2 ways and fail in n_2 ways.

$$\therefore P(B) = \frac{m_2}{m_2 + n_2}$$

Now there are four possibilities

A and B both may happen, then the number of ways = $m_1 \cdot m_2$.

A may happen and B may fail, then the number of ways = $m_1 \cdot n_2$

A may fail and B may happen, then the number of ways = $n_1 \cdot m_2$

A and B both may fail, then the number of ways = $n_1 \cdot n_2$

Thus, the total number of ways = $m_1 m_2 + m_1 n_2 + n_1 m_2 + n_1 n_2 = (m_1 + n_1)(m_2 + n_2)$

Hence the probabilities of the happening of both A and B

$$P(AB) = \frac{m_1 m_2}{(m_1 + n_1)(m_2 + n_2)} = \frac{m_1}{m_1 + n_1} \cdot \frac{m_2}{m_2 + n_2} = P(A) \cdot P(B)$$

Proved.

Example 6. An article manufactured by a company consists of two parts A and B. In the process of manufacture of part A, 9 out of 100 are likely to be defective. Similarly, 5 out of 100 are likely to be defective in the manufacture of part B. Calculate the probability that the assembled article will not be defective (assuming that the events of finding the part A non-defective and that of B are independent).

Solution. Probability that part A will be defective = $\frac{9}{100}$

$$\text{Probability that part A will not be defective} = \left(1 - \frac{9}{100}\right) = \frac{91}{100}$$

$$\text{Probability that part B will be defective} = \frac{5}{100}$$

$$\text{Probability that part B will not be defective} = \left(1 - \frac{5}{100}\right) = \frac{95}{100}$$

Probability that the assembled article will not be defective = (Probability that part A will not be defective) \times (Probability that part B will not be defective)

$$= \left(\frac{91}{100}\right) \times \left(\frac{95}{100}\right) = 0.8645 \quad \text{Ans.}$$

Example 7. The probability that machine A will be performing an usual function in 5 years'

time is $\frac{1}{4}$, while the probability that machine B will still be operating usefully at the

end of the same period, is $\frac{1}{3}$

Find the probability in the following cases that in 5 years time :

(i) Both machines will be performing an usual function.

(ii) Neither will be operating.

(iii) Only machine B will be operating.

(iv) At least one of the machines will be operating.

Solution. $P(A \text{ operating usefully}) = \frac{1}{4}$, so $q(A) = 1 - \frac{1}{4} = \frac{3}{4}$

$P(B \text{ operating usefully}) = \frac{1}{3}$, so $q(B) = 1 - \frac{1}{3} = \frac{2}{3}$

$$(i) P(\text{Both } A \text{ and } B \text{ will operate usefully}) = P(A) \cdot P(B) = \left(\frac{1}{4}\right) \times \left(\frac{1}{3}\right) = \frac{1}{12}$$

$$(ii) P(\text{Neither will be operating}) = q(A) \cdot q(B) = \left(\frac{3}{4}\right) \left(\frac{2}{3}\right) = \frac{1}{2}$$

$$(iii) P(\text{Only } B \text{ will be operating}) = P(B) \times q(A) = \left(\frac{1}{3}\right) \times \left(\frac{3}{4}\right) = \frac{1}{4}$$

$$(iv) P(\text{At least one of the machines will be operating}) \\ = 1 - P(\text{none of them operates})$$

$$= 1 - \frac{1}{2} = \frac{1}{2} \quad \text{Ans.}$$

Example 8. There are two groups of subjects one of which consists of 5 science and 3 engineering subjects and the other consists of 3 science and 5 engineering subjects. An unbiased

die is cast. If number 3 or number 5 turns up, a subject is selected at random from the first group, otherwise the subject is selected at random from the second group. Find the probability that an engineering subject is selected ultimately.

(A.M.I.E.T.E., Summer 2000)

Solution. Probability of turning up 3 or 5 = $\frac{2}{6} = \frac{1}{3}$

Probability of selecting engineering subject from first group = $\frac{3}{8}$

Now the probability of selecting engineering subject from first group on turning up 3 or 5

$$= \left(\frac{1}{3}\right) \times \left(\frac{3}{8}\right) = \frac{1}{8} \quad \dots (1)$$

Probability of not turning up 3 or 5 = $1 - \frac{1}{3} = \frac{2}{3}$

Probability of selecting engineering subject from second group = $\frac{5}{8}$

Now probability of selecting engineering subject from second group on not turning up 3 or 5

$$= \frac{2}{3} \times \frac{5}{8} = \frac{5}{12} \quad \dots (2)$$

Probability of the selection of engineering subject = $\frac{1}{8} + \frac{5}{12}$ [From (1) and (2)]

$$= \frac{13}{24} \quad \text{Ans.}$$

Example 9. An urn contains nine balls, two of which are red, three blue and four black. Three balls are drawn from the urn at random. What is the probability that

(i) the three balls are of different colours?

(ii) the three balls are of the same colour? (A.M.I.E., Summer 2000)

Solution.

Urn contains 2 Red balls, 3 Blue balls and 4 Black balls.

(i) Three balls will be of different colours if one ball is red, one blue and one black ball are drawn.

$$\text{Required probability} = \frac{{}^2C_1 \times {}^3C_1 \times {}^4C_1}{{}^9C_3} = \frac{2 \times 3 \times 4}{84} = \frac{2}{7} \quad \text{Ans.}$$

(ii) Three balls will be of the same colour if either 3 blue balls or 3 black balls are drawn.

$P(3 \text{ Blue balls or } 3 \text{ Black balls}) = P(3 \text{ Blue balls}) + P(3 \text{ Black balls})$

$$= \frac{{}^3C_3}{{}^9C_3} + \frac{{}^4C_3}{{}^9C_3} = \frac{1+4}{84} = \frac{5}{84} \quad \text{Ans.}$$

Example 10. A bag contains 10 white and 15 black balls. Two balls are drawn in succession. What is the probability that first is white and second is black?

Solution. Probability of drawing one white ball = $\frac{10}{25}$

Probability of drawing one black ball without replacement = $15/24$

Required probability of drawing first white ball and second black ball

$$= \frac{10}{25} \times \frac{15}{24} = \frac{1}{4} \quad \text{Ans.}$$

Example 11. A committee is to be formed by choosing two boys and four girls out of a group of five boys and six girls. What is the probability that a particular boy named A and a particular girl named B are selected in the committee?

Solution. Two boys are to be selected out of 5 boys. A particular boy A is to be included in the committee. It means that only 1 boy is to be selected out of 4 boys.

$$\text{Number of ways of selection} = {}^4C_1$$

Similarly a girl B is to be included in the committee.

Then only 3 girls are to be selected out of 5 girls.

$$\text{Number of ways of selection} = {}^5C_3$$

$$\text{Required probability} = \frac{{}^4C_1 \times {}^5C_3}{{}^5C_2 \times {}^6C_4} = \frac{4 \times 10}{10 \times 15} = \frac{4}{15} \quad \text{Ans.}$$

Example 12. Three groups of children contain respectively 3 girls and 1 boy; 2 girls and 2 boys; 1 girl and 3 boys. One child is selected at random from each group. Find the chance of selecting 1 girl and 2 boys.

Solution. There are three ways of selecting 1 girl and two boys.

I way : Girl is selected from first group, boy from second group and second boy from third group.

$$\text{Probability of the selection of (Girl + Boy + Boy)} = \frac{3}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{18}{64}$$

II way : Boy is selected from first group, girl from second group and second boy from third group.

$$\text{Probability of the selection of (Boy + Girl + Boy)} = \frac{1}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{6}{64}$$

III way : Boy is selected from first group, second boy from second group and the girl from the third group.

$$\text{Probability of selection of (Boy + Boy + Girl)} = \frac{1}{4} \times \frac{2}{4} \times \frac{1}{4} = \frac{2}{64}$$

$$\text{Total probability} = \frac{18}{64} + \frac{6}{64} + \frac{2}{64} = \frac{26}{64} = \frac{13}{32} \quad \text{Ans.}$$

Example 13. The number of children in a family in a region are either 0, 1 or 2 with probability 0.2, 0.3 and 0.5 respectively. The probability of each child being a boy or girl 0.5. Find the probability that a family has no boy.

Solution. Here there are three types of families

(i) Probability of zero child (boys) = 0.2

(ii)	Boy	Girl
	0	1
	1	0

Probability of zero boy in case II = $0.3 \times 0.5 = 0.15$

(iii)	Boy	Girl
	0	2
	1	1
	2	0

$$\text{In this case probability of zero boy} = 0.5 \times \frac{1}{3} = 0.167$$

Considering all the three cases, the probability of zero boy

$$= 0.2 + 0.15 + 0.167 = 0.517$$

Ans.

Example 14. A husband and wife appear in an interview for two vacancies in the same

post. The probability of husband's selection is $\frac{1}{7}$ and that of wife's selection is

$\frac{1}{5}$. What is the probability that

- (i) both of them will be selected.
- (ii) only one of them will be selected, and
- (iii) none of them will be selected?

Solution. $P(\text{husband's selection}) = \frac{1}{7}$, $P(\text{wife's selection}) = \frac{1}{5}$

$$(i) P(\text{both selected}) = \frac{1}{7} \times \frac{1}{5} = \frac{1}{35}$$

$$(ii) P(\text{only one selected}) = P(\text{only husband's selection}) + P(\text{only wife's selection})$$

$$= \frac{1}{7} \times \frac{4}{5} + \frac{1}{5} \times \frac{6}{7} = \frac{10}{35} = \frac{2}{7}$$

$$(iii) P(\text{none of them will be selected}) = \frac{6}{7} \times \frac{4}{5} = \frac{24}{35}$$

Ans.

Example 15. A problem of statistics is given to three students A, B and C whose chances of solving it are $\frac{1}{2}$, $\frac{3}{4}$ and $\frac{1}{4}$ respectively. What is the probability that the problem will be solved?

Solution. The probability that A can solve the problem = $\frac{1}{2}$

The probability that A cannot solve the problem = $1 - \frac{1}{2}$.

Similarly the probability that B and C cannot solve the problem are

$$\left(1 - \frac{3}{4}\right) \text{ and } \left(1 - \frac{1}{4}\right)$$

\therefore The probability that A, B, C cannot solve the problem

$$= \left(1 - \frac{1}{2}\right) \times \left(1 - \frac{3}{4}\right) \times \left(1 - \frac{1}{4}\right) = \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4} = \frac{3}{32}$$

Hence, the probability that the problem can be solved = $1 - \frac{3}{32} = \frac{29}{32}$ **Ans.**

Example 16. A student takes his examination in four subjects α , β , γ , δ . He estimates his

chances of passing in α as $\frac{4}{5}$, in β as $\frac{3}{4}$, in γ as $\frac{5}{6}$ and in δ as $\frac{2}{3}$. To qualify, he must pass in α and at least two other subjects. What is the probability that he qualifies? (AMIETE, Dec. 2010)

Solution. $P(\alpha) = \frac{4}{5}$, $P(\beta) = \frac{3}{4}$, $P(\gamma) = \frac{5}{6}$, $P(\delta) = \frac{2}{3}$

There are four possibilities of passing at least two subjects.

$$(i) \text{Probability of passing } \beta, \gamma \text{ and failing } \delta = \frac{3}{4} \times \frac{5}{6} \times \left(1 - \frac{2}{3}\right) = \frac{3}{4} \times \frac{5}{6} \times \frac{1}{3} = \frac{5}{24}$$

$$(ii) \text{Probability of passing } \gamma, \delta \text{ and failing } \beta = \frac{5}{6} \times \frac{2}{3} \times \left(1 - \frac{3}{4}\right) = \frac{5}{6} \times \frac{2}{3} \times \frac{1}{4} = \frac{5}{36}$$

$$(iii) \text{Probability of passing } \delta, \beta \text{ and failing } \gamma = \frac{2}{3} \times \frac{3}{4} \times \left(1 - \frac{5}{6}\right) = \frac{2}{3} \times \frac{3}{4} \times \frac{1}{6} = \frac{1}{12}$$

$$(iv) \text{ Probability of passing } \beta, \gamma, \delta = \frac{3}{4} \times \frac{5}{6} \times \frac{2}{3} = \frac{5}{12}$$

$$\text{Probability of passing at least two subjects} = \frac{5}{24} + \frac{5}{36} + \frac{1}{12} + \frac{5}{12} = \frac{61}{72}$$

$$\text{Probability of passing } \alpha \text{ and at least two subjects.} = \frac{4}{5} \times \frac{61}{72} = \frac{61}{90} \quad \text{Ans.}$$

Example 17. There are 6 positive and 8 negative numbers. Four numbers are chosen at random, without replacement, and multiplied. What is the probability that the product is a positive number?

Solution. To get from the product of four numbers, a positive number, the possible combinations are as follows :

S. No.	Out of 6 Positive Numbers	Out of 8 Negative Numbers	Positive Numbers
1.	4	0	${}^6C_4 \times {}^8C_0 = \frac{6 \times 5}{1 \times 2} \times 1 = 15$
2.	2	2	${}^6C_2 \times {}^8C_2 = \frac{6 \times 5}{1 \times 2} \times \frac{8 \times 7}{1 \times 2} = 420$
3.	0	4	${}^6C_0 \times {}^8C_4 = 1 \times \frac{8 \times 7 \times 6 \times 5}{1 \times 2 \times 3 \times 4} = 70$
Total = 505			

$$\text{Probability} = \frac{{}^6C_4 \times {}^8C_0 + {}^6C_2 \times {}^8C_2 + {}^6C_0 \times {}^8C_4}{{}^{14}C_4} = \frac{15 + 420 + 70}{14 \times 13 \times 12 \times 11} = \frac{505 \times 4 \times 3 \times 2 \times 1}{14 \times 13 \times 12 \times 11} = \frac{505}{1001} \quad \text{Ans.}$$

$$1 \times 2 \times 3 \times 4$$

Example 18. A can hit a target 3 times in 5 shots, B 2 times in 5 shots and C three times in 4 shots. All of them fire one shot each simultaneously at the target. What is the probability that

- (i) 2 shots hit (ii) At least two shots hit?

Solution. Probability of A hitting the target = $\frac{3}{5}$

$$\text{Probability of B hitting the target} = \frac{2}{5}$$

$$\text{Probability of C hitting the target} = \frac{3}{4}$$

Probability that 2 shots hit the target

$$= P(A)P(B)q(C) + P(A)P(C)q(B) + P(B)P(C)q(A)$$

$$= \frac{3}{5} \times \frac{2}{5} \times \left(1 - \frac{3}{4}\right) + \frac{3}{5} \times \frac{3}{4} \times \left(1 - \frac{2}{5}\right) + \frac{2}{5} \times \frac{3}{4} \times \left(1 - \frac{3}{5}\right) = \frac{6}{25} \times \frac{1}{4} + \frac{9}{20} \times \frac{3}{5} + \frac{6}{20} \times \frac{2}{5}$$

$$= \frac{6+27+12}{100} = \frac{45}{100} = \frac{9}{20} \quad \text{Ans.}$$

(ii) Probability of at least two shots hitting the target

$$= \text{Probability of 2 shots} + \text{probability of 3 shots hitting the target}$$

$$= \frac{9}{20} + P(A)P(B)P(C) = \frac{9}{20} + \frac{3}{5} \times \frac{2}{5} \times \frac{3}{4} = \frac{63}{100} \quad \text{Ans.}$$

Example 19. *A and B take turns in throwing two dice, the first to throw 10 being awarded the prize. Show that if A has the first throw, their chances of winning are in the ratio 12:11. (AMIETE, Dec. 2009)*

Solution. The combinations of throwing 10 from two dice can be

$$(6+4), (4+6), (5+5).$$

The number of combinations is 3.

Total combinations from two dice = $6 \times 6 = 36$.

$$\therefore \text{The probability of throwing 10} = p = \frac{3}{36} = \frac{1}{12}$$

$$\text{The probability of not getting 10} = q = 1 - \left(\frac{1}{12}\right) = \frac{11}{12}$$

If A is to win, he should throw 10 in either the first, the third, the fifth, ... throws.

$$\text{Their respective probabilities are} = p, q^2 p, q^4 p, \dots = \frac{1}{12}, \left(\frac{11}{12}\right)^2 \frac{1}{12}, \left(\frac{11}{12}\right)^4 \frac{1}{12}, \dots$$

$$\begin{aligned} A's \text{ total probability of winning} &= \frac{1}{12} + \left(\frac{11}{12}\right)^2 \cdot \frac{1}{12} + \left(\frac{11}{12}\right)^4 \cdot \frac{1}{12} + \dots \\ &= \frac{\frac{1}{12}}{1 - \left(\frac{11}{12}\right)^2} = \frac{12}{23} \quad \left[\text{This is infinite G.P. Its sum} = \frac{a}{1-r} \right] \end{aligned}$$

B can win in either 2nd, 4th, 6th ... throws.

So B's total chance of winning = $qp + q^3 p + q^5 p + \dots$

$$= \left(\frac{11}{12}\right) \left(\frac{1}{12}\right) + \left(\frac{11}{12}\right)^3 \left(\frac{1}{12}\right) + \left(\frac{11}{12}\right)^5 \left(\frac{1}{12}\right) + \dots = \frac{\left(\frac{11}{12}\right) \left(\frac{1}{12}\right)}{1 - \left(\frac{11}{12}\right)^2} = \frac{11}{23}$$

$$\text{Hence } A's \text{ chance to } B's \text{ chance} = \frac{12}{23} : \frac{11}{23} = 12 : 11$$

Proved.

Example 20. *A and B throw alternatively a pair of dice. A wins if he throws 6 before B throws 7 and B wins if he throws 7 before A throws 6. Find their respective chances of winning, if A begins.*

Solution. Number of ways of throwing 6

$$\text{i.e., } (1+5), (2+4), (3+3), (4+2), (5+1) = 5.$$

$$\text{Probability of throwing 6} = \frac{5}{36} = p_1, \quad q_1 = \frac{31}{36}$$

Number of ways of throwing 7

$$\text{i.e., } (1+6), (2+5), (3+4), (4+3), (5+2), (6+1) = 6$$

$$\text{Probability of throwing 6} = \frac{6}{36} = \frac{1}{6} = P_2, \quad q_2 = \frac{5}{6}$$

$$P(A) = p_1 + q_1 q_2 p_1 + q_1^2 q_2^2 p_1 + \dots$$

$$P(B) = q_1 p_2 + q_1^2 q_2 p_2 + q_1^3 q_2^2 p_2 + \dots$$

$$\text{Probability of A's winning} = p_1 + q_1 q_2 p_1 + q_1^2 q_2^2 p_1 + \dots$$

$$= \frac{p_1}{1-q_1q_2} = \frac{\frac{5}{36}}{1-\frac{31}{36} \times \frac{5}{6}} = \frac{5}{36} \times \frac{36 \times 6}{61} = \frac{30}{61}$$

Probability of B's winning = $q_1 p_2 + q_1^2 q_2 p_2 + q_1^3 q_2^2 p_2 + \dots$

$$= \frac{q_1 p_2}{1-q_1q_2} = \frac{\frac{31}{36} \times \frac{1}{6}}{1-\left(\frac{31}{36}\right)\left(\frac{5}{6}\right)} = \frac{31}{36 \times 6} \times \frac{36 \times 6}{61} = \frac{31}{61}$$

Ans.

EXERCISE 61.2

1. The probability that Nirmal will solve a problem is $\frac{2}{3}$ and the probability that Satyajit will solve it is $\frac{3}{4}$. What is the probability that (a) the problem will be solved (b) neither can solve it.

Ans. (a) $\frac{11}{12}$, (b) $\frac{1}{12}$

2. An urn contains 13 balls numbering 1 to 13. Find the probability that a ball selected at random is a ball with number that is a multiple of 3 or 4.

Ans. $\frac{6}{13}$

3. Four persons are chosen at random from a group containing 3 men, 2 women, and 4 children. Show that the probability that exactly two of them will be children is $\frac{10}{21}$.

4. A five digit number is formed by using the digits 0, 1, 2, 3, 4 and 5 without repetition. Find the probability that the number is divisible by 6.

Ans. $\frac{4}{25}$

5. The chances that doctor A will diagnose a disease X correctly is 60%. The chances that a patient will die by his treatment after correct diagnosis is 40% and the chances of death by wrong diagnosis is 70%. A patient of doctor A, who had disease X, died, what is the chance that his disease was diagnosed correctly.

Ans. $\frac{6}{13}$

6. An anti-aircraft gun can take a maximum of four shots on enemy's plane moving from it. The probabilities of hitting the plane at first, second, third and fourth shots are 0.4, 0.3, 0.2 and 0.1 respectively. Find the probability that the gun hits the plane.

Ans. 0.6976.

7. An electronic component consists of three parts. Each part has probability 0.99 of performing satisfactorily. The component fails if two or more parts do not perform satisfactorily. Assuming that the parts perform independently, determine the probability that the component does not perform satisfactorily.

Ans. 0.000298

8. The face cards are removed from a full pack. Out of the remaining 40 cards, 4 are drawn at random. What is the probability that they belong to different suits?

Ans. $\frac{1000}{9139}$

9. Of the cigarette smoking population, 70% are men and 30% women, 10% of these men and 20% of these women smoke 'WILLS.' What is the probability that a person seen smoking a 'WILLS' will be a man.

Ans. $\frac{7}{13}$

10. A machine contains a component C that is vital to its operation. The reliability of component C is 80%. To improve the reliability of a machine, a similar component is used in parallel to form a system S. The machine will work provided that one of these components functions correctly. Calculate the reliability of the system S.

Ans. 96%

11. The odds that a book will be favourably reviewed by three independent critics are 5 to 2, 4 to 3, 3 to 4 respectively. What is the probability that of the three reviews, a majority will be favourable?

$$\text{Ans. } \frac{209}{343}$$

12. A man takes a step forward with probability 0.4 and backward with probability 0.6. Find the probability that at the end of 11 steps, he is just one step away from the starting point.

$$\text{Ans. } 0.5263$$

13. A candidate is selected for interview for three posts. For the first post there are three candidates, for the second there are 4, and for the third are 2. What is the chance of getting at least one post?

$$(A.M.I.E., Summer 2001) \text{ Ans. } \frac{3}{4}$$

14. The chance of hitting a target by a bomb is 50% when 4 bombs are dropped, what is the probability of destroying the target, if one bomb is just sufficient to destroy it. *(A.M.I.E., Winter 2003)*

$$\text{Ans. } \frac{15}{16}$$

15. Tick \checkmark the correct answer :

- (i) A, B, C in order toss a coin, the first to throw a head wins. Assuming if A begins and the game continues indefinitely their respective chances of winning are:

$$(a) \frac{4}{7}, \frac{2}{7}, \frac{1}{7} \quad (b) \frac{1}{7}, \frac{4}{7}, \frac{2}{7} \quad (c) \frac{2}{7}, \frac{4}{7}, \frac{1}{7} \quad (d) \text{None of these}$$

$$(A.M.I.E., winter 2000) \text{ Ans. (a)}$$

- (ii) An unbiased die with faces marked 1, 2, 3, 4, 5, 6 is rolled 4 times, out of four face values obtained, the probability that the minimum face value is not less than 2 and the maximum face value is not greater than 5 is then

$$(a) \frac{16}{81} \quad (b) \frac{2}{9} \quad (c) \frac{80}{81} \quad (d) \frac{8}{9}$$

$$(A.M.I.E.T.E., Summer 2000) \text{ Ans. (a)}$$

- (iii) India plays two matches each with England and Australia. In any match the probability of its getting points 0, 1 and 2 are 0.45, 0.05 and 0.5 respectively. Assuming the outcomes are independent, the probability that India gets at least seven points is

$$(a) 0.8750 \quad (b) 0.0875 \quad (c) 0.0625 \quad (d) 0.0250$$

$$(A.M.I.E.T.E., Summer 2001) \text{ Ans. (b)}$$

Fill up the blanks:

- (iv) Probability of any event can not be greater than _____ and less than _____.

$$(A.M.I.E.T.E., Winter 2001) \text{ Ans. } 0, 1$$

61.5 CONDITIONAL PROBABILITY

Let A and B be two events of a sample space S and let $P(B) \neq 0$. Then conditional probability of the event A , given B , denoted by $P(A/B)$, is defined by

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \dots (1)$$

Theorem. If the events A and B defined on a sample space S of a random experiment are independent, then

$$P(A/B) = P(A) \text{ and } P(B/A) = P(B)$$

Proof. A and B are given to be independent events,

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

$$\Rightarrow P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

$$\Rightarrow P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B) \cdot P(A)}{P(A)} = P(B)$$

61.6 BAYE'S THEOREM

If $B_1, B_2, B_3, \dots, B_n$ are mutually exclusive events with $P(B_i) \neq 0$, ($i = 1, 2, \dots, n$) of a random experiment then for any arbitrary event A of the sample space of the above experiment with $P(A) > 0$, we have

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{\sum_{i=1}^n P(B_i)P(A/B_i)} \quad (\text{for } n=3)$$

$$P(B_2/A) = \frac{P(B_2)P(A/B_2)}{P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + P(B_3)P(A/B_3)}$$

Proof. Let S be the sample space of the random experiment.

The events B_1, B_2, \dots, B_n being exhaustive

$$S = B_1 \cup B_2 \cup \dots \cup B_n$$

$$[\because A \subset S]$$

$$\begin{aligned} \therefore A = A \cap S &= A \cap (B_1 \cup B_2 \cup \dots \cup B_n) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n) \quad [\text{Distributive Law}] \end{aligned}$$

$$\begin{aligned} \Rightarrow P(A) &= P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_n) \\ &= P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + \dots + P(B_n)P(A/B_n) \\ &= \sum_{i=1}^n P(B_i)P(A/B_i) \end{aligned} \quad \dots (1)$$

$$\text{Now, } P(A \cap B_i) = P(A)P(B_i/A)$$

$$\Rightarrow P(B_i/A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(B_i)P(A/B_i)}{\sum_{i=1}^n P(B_i)P(A/B_i)} \quad [\text{Using (1)}]$$

Note. $P(B)$ is the probability of occurrence B . If we are told that the event A has already occurred.

On knowing about the event A , $P(B)$ is changed to $P(B/A)$. With the help of Baye's theorem we can calculate $P(B/A)$.

Example 21. An urn I contains 3 white and 4 red balls and an urn II contains 5 white and 6 red balls. One ball is drawn at random from one of the urns and is found to be white. Find the probability that it was drawn from urn I.

Solution. Let U_1 : the ball is drawn from urn I

U_2 : the ball is drawn from urn II

W : the ball is white.

We have to find $P(U_1/W)$

By Baye's Theorem

$$P(U_1/W) = \frac{P(U_1)P(W/U_1)}{P(U_1)P(W/U_1) + P(U_2)P(W/U_2)} \quad \dots (1)$$

Since two urns are equally likely to be selected, $P(U_1) = P(U_2) = \frac{1}{2}$

$$P(W/U_1) = P(\text{a white ball is drawn from urn I}) = \frac{3}{7}$$

$$P(W/U_2) = P(\text{a white ball is drawn from urn II}) = \frac{5}{11}$$

$$\therefore \text{From (1), } P(U_1/W) = \frac{\frac{1}{2} \times \frac{3}{7}}{\frac{1}{2} \times \frac{3}{7} + \frac{1}{2} \times \frac{5}{11}} = \frac{33}{68}$$

Ans.

Example 22. Three urns contains 6 red, 4 black; 4 red, 6 black; 5 red, 5 black balls respectively. One of the urns is selected at random and a ball is drawn from it. If the ball drawn is red find the probability that it is drawn from the first urn.

Solution. Let U_1 : the ball is drawn from U_1 ,

U_2 : the ball is drawn from U_2 ,

U_3 : the ball is drawn from U_3 ,

R : the ball is red.

We have to find $P(U_1/R)$.

By Baye's Theorem,

$$P(U_1/R) = \frac{P(U_1)P(R/U_1)}{P(U_1)P(R/U_1) + P(U_2)P(R/U_2) + P(U_3)P(R/U_3)} \quad \dots (1)$$

Since the three urns are equally likely to be selected $P(U_1) = P(U_2) = P(U_3) = \frac{1}{3}$

$$\text{Also } P(R/U_1) = P(\text{a red ball is drawn from urn I}) = \frac{6}{10}$$

$$P(R/U_2) = P(\text{a red ball is drawn from urn II}) = \frac{4}{10}$$

$$P(R/U_3) = P(\text{a red ball is drawn from urn III}) = \frac{5}{10}$$

$$\therefore \text{From (1), we have } P(U_1/R) = \frac{\frac{1}{3} \times \frac{6}{10}}{\frac{1}{3} \times \frac{6}{10} + \frac{1}{3} \times \frac{4}{10} + \frac{1}{3} \times \frac{5}{10}} = \frac{2}{5}$$

Ans.

Example 23. In a bolt factory, machines A, B and C manufacture respectively 25%, 35% and 40% of the total. If their output 5, 4 and 2 per cent are defective bolts. A bolt is drawn at random from the product and is found to be defective. What is the probability that it was manufactured by machine B?

Solution. A : bolt is manufactured by machine A.

B : bolt is manufactured by machine B.

C : bolt is manufactured by machine C.

$$P(A) = 0.25, P(B) = 0.35, P(C) = 0.40$$

The probability of drawing a defective bolt manufactured by machine A is $P(D/A) = 0.05$

Similarly, $P(D/B) = 0.04$ and $P(D/C) = 0.02$

By Baye's theorem

$$\begin{aligned} P(B/D) &= \frac{P(B)P(D/B)}{P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C)} \\ &= \frac{0.35 \times 0.04}{0.25 \times 0.05 + 0.35 \times 0.04 + 0.40 \times 0.02} = 0.41 \end{aligned} \quad \text{Ans.}$$

Example 24. An insurance company insured 2000 scooter drivers 4000 car drivers and 6000 truck drivers. The probability of accidents are 0.01, 0.03 and 0.15 respectively. One of the insured persons meets with an accident. What is the probability that he is a scooter driver?

(AMIETE, Dec. 2009)

Solution. Let E_1, E_2, E_3 and A be the events defined as follows :

E_1 = person chosen is a scooter driver

E_2 = person chosen is a car driver

E_3 = person chosen is a truck driver, and

A = person meets with an accident.

We have,

$$n(E_1) = 2000, n(E_2) = 4000, n(E_3) = 6000$$

$$\text{Total number of persons} = 2000 + 4000 + 6000 = 12000.$$

Therefore,

$$P(E_1) = \frac{2000}{12000} = \frac{1}{6} \text{ and } P(E_2) = \frac{4000}{12000} = \frac{1}{3} \text{ and } P(E_3) = \frac{6000}{12000} = \frac{1}{2}$$

It is given that

$$\begin{aligned} P(A/E_1) &= \text{Probability that a person meets with an accident given that he is a scooter driver} \\ &= 0.01. \end{aligned}$$

$$\text{Similarly, } P(A/E_2) = 0.03 \text{ and } P(A/E_3) = 0.15$$

We are required to find $P(E_1/A)$ i.e., given that the person meets with an accident, what is the probability that he was a scooter driver.

By Bayes' rule, we have

$$\begin{aligned} P(E_1/A) &= \frac{P(E_1)P(A/E_1)}{P(E_1)P(A/E_1) + P(E_2)P(A/E_2) + P(E_3)P(A/E_3)} \\ \Rightarrow P(E_1/A) &= \frac{\frac{1}{6} \times 0.01}{\frac{1}{6} \times 0.01 + \frac{1}{3} \times 0.03 + \frac{1}{2} \times 0.15} = \frac{1}{1+6+45} = \frac{1}{52} \end{aligned} \quad \text{Ans.}$$

EXERCISE 61.3

1. A bag contains 3 coins of which one is two headed and the other two are normal & fair. A coin is selected at random and tossed 4 times in succession. If all the four times it appears to be head what is the probability that the two headed coin was selected.

(AMIETE, JUNE 2010) Ans. $\frac{8}{9}$

2. An airline knows that 13% of the people who make reservations on a certain flight will not turn up. Consequently their policy is to sell 12 tickets which can accommodate only 10. What is the probability that everyone who turns up on a given day is accommodated? (AMIETE, JUNE 2010)

4.17 RANDOM VARIABLE

Let E be a random experiment and S be a sample space associated with it. A function X assigning to every element $s \in S$, a real number, $X(s)$ is called a random variable.

A random variable is any numerical quantity whose value will be determined by the outcome of a random experiment and which have a specific range and a definable probability associated with each value.

In other words, if the numerical values assumed by a variable are the result of some chance factors, so that a particular value cannot be exactly predicted in advance, the variable is then called a *random variable*. A random variable is also called '*chance variable*' or '*stochastic variable*'. Random variables are denoted by capital letters, usually, from the last part of the alphabet, for instance, X , Y , Z , etc.

4.17.1 Definition

A real valued function, X , defined on a sample space S of a random experiment, E , is called a random variable which assigns to each sample point (or to every outcome of the random experiment E) $s \in S$ one and only one real number $X(s) = x$. The domain of the random variable X is the sample space S and its range, A is the non empty set of real numbers such that

1. The set $\{X(s) \leq x\}$ is an event for any real number x .
2. There corresponds a well defined unique probability of the events, "X assumes the value x " and "X assumes any value in the interval".

e.g., If three coins are tossed, then the sample space contains 8 sample points. Let the random variable X denote "the number of heads". Then X is a real valued function over S with a space A which has four elements 0, 1, 2, 3, 4 as explained below:

Sample Point	$X(s)$	$X = \text{no. of heads}$	Probability
TTT	$X(\text{TTT})$	0	$\frac{1}{8}$
HTT, THT, TTH	$X(\text{HTT}), X(\text{THT}), X(\text{TTH})$	1	$\frac{3}{8}$
THH, HTH, HHT	$X(\text{THH}), X(\text{HTH}), X(\text{HHT})$	2	$\frac{3}{8}$
HHH	$X(\text{HHH})$	3	$\frac{1}{8}$

Thus A the space of X is a set of real numbers given by

$$A = \{x : x = X(s), s \in S\}.$$

In this example, we have

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

and

$$A = \{0, 1, 2, 3\}.$$

If S has elements which are themselves real numbers then $X(s) = s$ and $A = S$.

As in case of rolling a die $S = \{1, 2, 3, 4, 5, 6\} = A$.

4.18 TYPES OF RANDOM VARIABLES

There are two types of random variables:

(1) **Discrete Random Variable.** A *discrete random variable* is one which can assume only isolated values. For example,

(i) the number of heads in 4 tosses of a coin is a discrete random variable as it cannot assume values other than 0, 1, 2, 3, 4.

(ii) the number of aces in a draw of 2 cards from a well shuffled deck is a random variable as it can take the values 0, 1, 2 only.

(2) **Continuous Random Variable.** A *continuous random variable* is one which can assume any value within an interval, i.e. all values of a continuous scale. For example (i) the weights (in kg) of a group of individuals, (ii) the heights of a group of individuals.

4.19 PROBABILITY FUNCTION (OR PROBABILITY MASS FUNCTION)

Let x_1, x_2, x_3, \dots be the values of a discrete random variable X and let $p_1, p_2, p_3, \dots, p_i > 0, i = 1, 2, 3, \dots$ be the corresponding probabilities. That is $P(X = x_i) = p(x_i)$ or simply p_i .

A function $f(x)$ or $p(x)$ defined by

$$P(X = x) = p(x) = \begin{cases} p(x_i) \text{ or } p_i \text{ where } x = x_i, i = 1, 2, \dots \\ 0 \text{ otherwise} \end{cases}$$

is called the probability function of the (discrete) random variable X .

The probability function $p(x)$ yields the probability that the random variable X assumes any particular value x in its range.

The probability function (or often called frequency function) of the random variable X satisfies the following conditions:

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

$$(ii) \sum p(x_i) = 1.$$

Example. A random variable X takes values 1, 2, 3, ... with probability mass function

$\frac{\lambda^r}{r!}, r = 1, 2, 3, \dots \infty$. Find the value of λ .

Sol. $\sum_{r=1}^{\infty} \frac{\lambda^r}{r!}$ should be 1. Thus,

$$\sum_{r=1}^{\infty} \frac{\lambda^r}{r!} = 1$$

$$\Rightarrow \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots = 1$$

$$\Rightarrow e^\lambda - 1 = 1$$

$$\Rightarrow e^\lambda = 2$$

$$\lambda = \log_e 2.$$

4.20 PROBABILITY DISTRIBUTION OF A DISCRETE RANDOM VARIABLE

A table or a formula listing all possible values that a random variable can take on together with the respective probabilities, is called a probability distribution of the random variable.

In other words, the set of ordered pairs $[x_i, p(x_i)]$ is called the probability distribution of a discrete random variable X provided $p(x_i) \geq 0$ and $\sum p(x_i) = 1$.

e.g. (i) Let X = The number of points appearing in a throw of a die. Then the probability distribution of X is given by the probability function

$$p(x) = \begin{cases} \frac{1}{6}, & x = 1, 2, 3, 4, 5, 6 \\ 0, & \text{otherwise.} \end{cases}$$

This probability distribution may also be expressed as follows:

X:	1	2	3	4	5	6
----	---	---	---	---	---	---

p(x) = P(X = x):	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
------------------	---------------	---------------	---------------	---------------	---------------	---------------

(ii) Suppose a coin is tossed three times. Then the distribution of the number of heads is

X = x:	0	1	2	3
--------	---	---	---	---

P(X = x) = p(x):	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
------------------	---------------	---------------	---------------	---------------

Note. In general, both terms probability function and probability distribution are used interchangeably.

4.21 MEAN AND VARIANCE OF RANDOM VARIABLES

Let

$$\begin{array}{lll} X : x_1, & x_2, & x_3, \dots, x_n \\ P(X) : p_1, & p_2, & p_3, \dots, p_n \end{array}$$

be a discrete probability distribution.

We denote the *mean* by μ and define $\mu = \frac{\sum p_i x_i}{\sum p_i} = \sum p_i x_i$ ($\because \sum p_i = 1$)

Other names for the mean are *average* or *expected value* $E(X)$.

We denote the *variance* by σ^2 and define $\sigma^2 = \sum p_i(x_i - \mu)^2$

If μ is not a whole number, then $\sigma^2 = \sum p_i x_i^2 - \mu^2$

Standard deviation $\sigma = + \sqrt{\text{Variance}}$.

ILLUSTRATIVE EXAMPLES

Example 1. Five defective bulbs are accidentally mixed with twenty good ones. It is not possible to just look at a bulb and tell whether or not it is defective. Find the probability distribution of the number of defective bulbs, if four bulbs are drawn at random from this lot.

Sol. Let X denote the number of defective bulbs in 4. Clearly X can take the values 0, 1, 2, 3 or 4.

$$\text{Number of defective bulbs} = 5$$

$$\text{Number of good bulbs} = 20$$

$$\text{Total number of bulbs} = 25$$

$$P(X = 0) = P(\text{no defective}) = P(\text{all 4 good ones})$$

$$= \frac{^{20}C_4}{^{25}C_4} = \frac{20 \times 19 \times 18 \times 17}{25 \times 24 \times 23 \times 22} = \frac{969}{2530}$$

$$P(X = 1) = P(\text{one defective and 3 good ones}) = \frac{^5C_1 \times ^{20}C_3}{^{25}C_4} = \frac{1140}{2530}$$

$$P(X = 2) = P(\text{2 defectives and 2 good ones}) = \frac{^5C_2 \times ^{20}C_2}{^{25}C_4} = \frac{380}{2530}$$

$$P(X = 3) = P(\text{3 defectives and 1 good one}) = \frac{^5C_3 \times ^{20}C_1}{^{25}C_4} = \frac{40}{2530}$$

$$P(X = 4) = P(\text{all 4 defectives}) = \frac{^5C_4}{^{25}C_4} = \frac{1}{2530}$$

\therefore The probability distribution of the random variable X is

X:	0	1	2	3	4
P(X):	$\frac{969}{2530}$	$\frac{1140}{2530}$	$\frac{380}{2530}$	$\frac{40}{2530}$	$\frac{1}{2530}$

Example 2. A die is tossed thrice. A success is 'getting 1 or 6' on a toss. Find the mean and the variance of the number of successes.

Sol. Let X denote the number of success. Clearly X can take the values 0, 1, 2 or 3.

$$\text{Probability of success} = \frac{2}{6} = \frac{1}{3}; \quad \text{Probability of failure} = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P(X = 0) = P(\text{no success}) = P(\text{all 3 failures}) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}$$

$$P(X = 1) = P(\text{one success and 2 failures}) = {}^3C_1 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{12}{27}$$

$$P(X=2) = P(\text{two successes and one failure}) = {}^3C_2 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{6}{27}$$

$$P(X=3) = P(\text{all 3 successes}) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$$

\therefore The probability distribution of the random variable X is

X:	0	1	2	3
P(X):	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

To find the mean and variance

x_i	p_i	$p_i x_i$	$p_i x_i^2$
0	$\frac{8}{27}$	0	0
1	$\frac{12}{27}$	$\frac{12}{27}$	$\frac{12}{27}$
2	$\frac{6}{27}$	$\frac{12}{27}$	$\frac{24}{27}$
3	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{9}{27}$
		1	$\frac{5}{3}$

$$\text{Mean } \mu = \sum p_i x_i = 1$$

$$\text{Variance } \sigma^2 = \sum p_i x_i^2 - \mu^2 = \frac{5}{3} - 1 = \frac{2}{3}.$$

Example 3. A random variable X has the following probability function:

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

(i) Find k ,

(ii) Evaluate $P(X < 6)$, $P(X \geq 6)$, $P(3 < X \leq 6)$

(iii) Find the minimum value of x so that $P(X \leq x) > \frac{1}{2}$.

Sol. (i) Since $\sum_{x=0}^7 p(x) = 1$, we have

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1 \\ \Rightarrow 10k^2 + 9k - 1 = 0 \quad \Rightarrow \quad (10k - 1)(k + 1) = 0$$

$$\Rightarrow k = \frac{1}{10} \quad [\because p(x) \geq 0]$$

$$(ii) P(X < 6) = P(X = 0) + P(X = 1) + \dots + P(X = 5)$$

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

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

$$P(3 < X \leq 6) = P(X = 4) + P(X = 5) + P(X = 6) = 3k + k^2 + 2k^2 = \frac{3}{10} + \frac{3}{100} = \frac{33}{100}$$

$$(iii) P(X \leq 1) = k = \frac{1}{10} < \frac{1}{2};$$

$$P(X \leq 2) = k + 2k = \frac{3}{10} < \frac{1}{2}$$

$$P(X \leq 3) = k + 2k + 2k = \frac{5}{10} = \frac{1}{2};$$

$$P(X \leq 4) = k + 2k + 2k + 3k = \frac{8}{10} > \frac{1}{2}$$

\therefore The maximum value of x so that $P(X \leq x) > \frac{1}{2}$ is 4.

TEST YOUR KNOWLEDGE

1. Find the probability distribution of the number of doublets in four throws of a pair of dice.
2. Two bad eggs are mixed accidentally with 10 good ones. Find the probability distribution of the number of bad eggs in 3, drawn at random, without replacement, from this lot.
3. A die is tossed twice. Getting a number greater than 4 is considered a success. Find the variance of the probability distribution of the number of successes.
4. Two cards are drawn simultaneously from a well-shuffled deck of 52 cards. Compute the variance for the number of aces.
5. An urn contains 4 white and 3 red balls. Three balls are drawn, with replacement, from this urn. Find μ , σ^2 and σ for the number of red balls drawn.
6. A random variable X has the following probability distribution:

Values of X , x :	0	1	2	3	4	5	6	7	8
$p(x)$:	a	$3a$	$5a$	$7a$	$9a$	$11a$	$13a$	$15a$	$17a$

- (i) Determine the value of a .
- (ii) Find $P(X < 3)$, $P(X \geq 3)$, $P(2 \leq X < 5)$
- (iii) What is the smallest value of x for which $P(X \leq x) > 0.5$?

7. Find the standard deviation for the following discrete distribution:

x :	8	12	16	20	24
$p(x)$:	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{12}$

Answers

1. X : 0 1 2 3 4
 $P(X)$: $\frac{625}{1296}$ $\frac{500}{1296}$ $\frac{150}{1296}$ $\frac{20}{1296}$ $\frac{1}{1296}$
2. X : 0 1 2
 $P(X)$: $\frac{12}{22}$ $\frac{9}{22}$ $\frac{1}{22}$
3. $\frac{4}{9}$
4. $\frac{400}{2873}$
5. $\frac{9}{7}, \frac{36}{49}, \frac{6}{7}$
6. (i) $a = \frac{1}{81}$ (ii) $\frac{1}{9}, \frac{8}{9}, \frac{7}{27}$ (iii) 5
7. $2\sqrt{5}$.

4.22 CONTINUOUS DISTRIBUTIONS

So far we have dealt with discrete distributions i.e., the distribution in which the variate takes a finite set of values. But when we deal with variates like temperature, heights and weights, we find that variates can take an infinite set of values in a given interval, say, $a \leq x \leq b$. Such variates are called continuous variates and their distributions are accordingly known as *continuous distributions*.

$$y = \frac{1}{2} \sin x \quad (0 \leq x \leq \pi)$$

is an example of a continuous distribution, as x can assume all values lying between 0 and π .

4.23 PROBABILITY DENSITY FUNCTION

Let the probability of the variate x falling in the infinitesimal interval $(x - \frac{1}{2}dx, x + \frac{1}{2}dx)$ be expressed in the form $f(x) dx$, where $f(x)$ is a continuous function of x . Then $f(x)$ is called the *probability density function* or simply *density function*. The continuous curve $y = f(x)$ is called the *probability density curve* or briefly the *probability curve*.

Symbolically it is expressed as

$$P(x - \frac{1}{2}dx \leq X \leq x + \frac{1}{2}dx) = f(x) dx.$$

The interval of the variate may be finite or infinite. A function defined only for a finite interval say $f(x) = \phi(x)$ when $a \leq x \leq b$ can be put in the following form

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

The density function possesses the following two properties:

(1) $f(x) \geq 0$ for every x , as negative probability has no meaning,

(2) $\int_{-\infty}^{\infty} f(x) dx = 1$, this corresponds to the fact that the probability of an event that is

sure to happen is equal to unity.

For a density function, $f(x)$, the probability that the variate X falls in any interval (p, q) is given by

$$P(p \leq X < q) = \int_p^q f(x) dx.$$

Note 1. Any positive function of a variate x can be changed to give a probability density function $f(x)$ by multiplying it by a constant which will make the total area under the curve $y = f(x)$ equal to unity. For example we know that

$$\int_0^2 x(2-x) dx = \frac{4}{3},$$

hence if we multiply both sides by $\frac{3}{4}$, we get

$$\int_0^2 \frac{3}{4}x(2-x) dx = 1.$$

Hence a probability density function can be formed as given by

$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{3x(2-x)}{4}, & 0 \leq x \leq 2 \\ 0, & x > 2. \end{cases}$$

Note 2. Whenever $f(x)$ is constant throughout its interval, the variable is said to have a rectangular distribution of same probability.

ILLUSTRATIVE EXAMPLES

Example 1. If the function $f(x)$ is defined by $f(x) = ce^{-x}$, $0 \leq x \leq \infty$, find the value of c which changes $f(x)$ to a probability density function.

Sol. In order that $f(x)$ may be density function, we should have

$$(a) f(x) \geq 0 \text{ for every } x$$

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

Since e^{-x} is always positive for the values of x lying between 0 and ∞ , the condition will be satisfied if $c \geq 0$.

The second condition will be satisfied if

$$\int_0^{\infty} ce^{-x} dx = 1$$

i.e., if

$$\left[-ce^{-x} \right]_0^{\infty} = 1$$

i.e., if

$$c = 1.$$

Example 2. If $f(x)$ has probability density cx^2 , $0 < x < 1$, determine c and find the probability that $\frac{1}{3} < x < \frac{1}{2}$ i.e., $P\left(\frac{1}{3} < x < \frac{1}{2}\right)$.

Sol. $f(x)$ will have a probability density if $\int_0^1 cx^2 dx = 1$

i.e.,

$$\left[\frac{1}{3} cx^3 \right]_0^1 = 1, \text{ i.e., } c = 3.$$

$$P\left(\frac{1}{3} < x < \frac{1}{2}\right) = \int_{\frac{1}{3}}^{\frac{1}{2}} 3x^2 dx = \left[x^3 \right]_{\frac{1}{3}}^{\frac{1}{2}} = \left(\frac{1}{8} - \frac{1}{27} \right) = \frac{19}{216}.$$

4.24 VARIOUS MEASURES FOR CONTINUOUS PROBABILITY DISTRIBUTIONS

Let $f_X(x)$ or $f(x)$ be the p.d.f. of a random variable X where X is defined from a to b . Then,

$$(i) \text{ Arithmetic Mean} = \int_a^b x f(x) dx \quad \dots(1)$$

$$(ii) \text{ Harmonic Mean.} \text{ Harmonic mean } H \text{ is given by: } \frac{1}{H} = \int_a^b \frac{1}{x} \cdot f(x) dx \quad \dots[1(a)]$$

(iii) **Geometric Mean.** Geometric mean G is given by: $\log G = \int_a^b \log x \cdot f(x) dx$...[1(b)]

$$(iv) v_r \text{ (about origin)} = \int_a^b x^r \cdot f(x) dx \quad \dots(2)$$

$$\mu'_r \text{ (about the point } x = A) = \int_a^b (x - A)^r \cdot f(x) dx \quad \dots[2(a)]$$

$$\text{and } \mu_r \text{ (about mean)} = \int_a^b (x - \text{mean})^r \cdot f(x) dx \quad \dots[2(b)]$$

In particular, from (1) and (2), we have

$$v_1 \text{ (about origin)} = \text{Mean} = \int_a^b x f(x) dx \quad \text{and} \quad v_2 = \int_a^b x^2 f(x) dx$$

$$\text{Hence } \mu_2 = (v_2 - v_1^2) = \int_a^b x^2 f(x) dx - \left(\int_a^b x f(x) dx \right)^2 \quad \dots[2(c)]$$

From (2), on putting $r = 3$ and 4 respectively, we get the values of v_3 and v_4 and consequently the moments about mean can be obtained and hence β_1 and β_2 can be computed.

(v) **Median.** Median is the point which divides the entire distribution in two equal parts. In case of continuous distribution, median is the point which divides the total area into two equal parts. Thus if M is the median, then

$$\int_a^M f(x) dx = \int_M^b f(x) dx = \frac{1}{2} \quad \dots(3)$$

$$\text{Thus solving } \int_a^M f(x) dx = \frac{1}{2} \text{ or } \int_M^b f(x) dx = \frac{1}{2} \quad \dots[3(a)]$$

for M , we get the value of median.

(vi) **Mean Deviation.** Mean deviation about the mean μ_1' is given by:

$$\text{M.D.} = \int_a^b |x - \text{mean}| f(x) dx \quad \dots(4)$$

In general, mean deviation about an average ' A ' is given by:

$$\text{M.D. about } 'A' = \int_a^b |x - A| f(x) dx \quad \dots[4(a)]$$

(vii) **Quartiles and Deciles.** Q_1 and Q_3 are given by the equations:

$$\int_a^{Q_1} f(x) dx = \frac{1}{4} \text{ and } \int_a^{Q_3} f(x) dx = \frac{3}{4} \quad \dots(5)$$

$$D_i, i^{\text{th}} \text{ decile is given by: } \int_a^{D_i} f(x) dx = \frac{i}{10}; i = 1, 2, \dots, 9 \quad \dots[5(a)]$$

(viii) **Mode.** Mode is the value of x for which $f(x)$ is maximum. Mode is thus the solution of $f'(x) = 0$ and $f''(x) < 0$, provided it lies in $[a, b]$[6]

4.25 CUMULATIVE DISTRIBUTION FUNCTION

If $F(x) = \int_{-\infty}^x f(x) dx = P(X \leq x)$, then the function $F(x)$ is the probability that the value of the variate X will be $\leq x$.

$$F(b) = P(X \leq b).$$

$$\text{and } F(b) - F(a) = \int_a^b f(x) dx = P(a \leq X \leq b).$$

$F(x)$ is called the cumulative distribution function of x or simply the *distribution function*. The cumulative distribution function has the following properties:

- (1) $F'(x) = f(x) \geq 0$, so that $F(x)$ is a non decreasing function. This means that $dF(x) = f(x) dx$. This is known as *probability differential of X*.
- (2) $F(-\infty) = 0$
- (3) $F(\infty) = 1$.
- (4) $F(x)$ is a continuous function of x on the right.

Example 3. The distribution function of a random variable X is given by

$$F(x) = \begin{cases} 1 - (1+x)e^{-x}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases}$$

find the corresponding density function of random variable X .

Sol. The probability density function $= f(x) = \frac{d}{dx} [F(x)]$

$$f(x) = -e^{-x} + (1+x)e^{-x}$$

$$= \begin{cases} xe^{-x}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0. \end{cases}$$

Example 4. A random variable x has the density function

$$f(x) = k \cdot \frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

Determine k and the distribution function.

Sol. It will be a density function if

$$\int_{-\infty}^{\infty} k \cdot \frac{1}{1+x^2} dx = 1 \text{ i.e., } k \cdot \pi = 1 \text{ giving } k = 1/\pi$$

$$F(x) = \int \frac{1}{\pi} \cdot \frac{1}{1+x^2} dx = \frac{1}{\pi} \tan^{-1} x + C$$

But $F(-\infty)$ should be zero for distribution function

$$\therefore \frac{1}{\pi} \left(-\frac{\pi}{2} \right) + C = 0 \text{ giving } C = \frac{1}{2}.$$

$$\therefore F(x) = \frac{1}{\pi} \tan^{-1} x + \frac{1}{2} \text{ for } -\infty < x < \infty.$$

Example 5. The diameter, say X , of an electric cable, is assumed to be a continuous random variable with p.d.f.: $f(x) = 6x(1-x)$, $0 \leq x \leq 1$

- (i) Check that the above is a p.d.f.,
- (ii) Obtain an expression for the c.d.f. of X ,

(iii) Compute $P\left(X \leq \frac{1}{2} \mid \frac{1}{3} \leq X \leq \frac{2}{3}\right)$, and

(iv) Determine the number k such that $P(X < k) = P(X > k)$.

Sol. (i) Since $\int_0^1 f(x) dx = \int_0^1 6x(1-x) dx = 6 \left| \frac{x^2}{2} - \frac{x^3}{3} \right|_0^1 = 1$, $f(x)$ is a p.d.f.

$$(ii) F(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ \int_0^x 6t(1-t) dt = (3x^2 - 2x^3), & 0 < x \leq 1 \\ 1, & \text{if } x > 1 \end{cases}$$

$$(iii) P\left(X \leq \frac{1}{2} \mid \frac{1}{3} \leq X \leq \frac{2}{3}\right) = \frac{P\left(\frac{1}{3} \leq X \leq \frac{1}{2}\right)}{P\left(\frac{1}{3} \leq X \leq \frac{2}{3}\right)} = \frac{\int_{1/3}^{1/2} 6x(1-x) dx}{\int_{1/3}^{2/3} 6x(1-x) dx} = \frac{11/54}{13/27} = \frac{11}{26}$$

(iv) We have $P(X < k) = P(X > k)$

$$\Rightarrow \int_0^k 6x(1-x) dx = \int_k^1 6x(1-x) dx$$

$$\text{or} \quad 3k^2 - 2k^3 = 3(1-k^2) - 2(1-k^3)$$

$$\Rightarrow 4k^3 - 6k^2 + 1 = 0 \Rightarrow k = \frac{1}{2}, \frac{1 \pm \sqrt{3}}{2}.$$

The only admissible value of k in the given range is $\frac{1}{2}$. Hence the value of k is $\frac{1}{2}$.

Example 6. Let X be a continuous random variable with p.d.f. given by:

$$f(x) = \begin{cases} kx, & 0 \leq x < 1 \\ k, & 1 \leq x < 2 \\ -kx + 3k, & 2 \leq x < 3 \\ 0, & \text{elsewhere} \end{cases}$$

(i) Determine the constant k , (ii) Determine $F(x)$, the c.d.f., and

(iii) If x_1 , x_2 and x_3 are three independent observations from X , what is the probability that exactly one of these three numbers is larger than 1.5?

Sol. (i) Since $f(x)$ is the p.d.f. of X , we have:

$$\int_0^3 f(x) dx = \int_0^1 f(x) dx + \int_1^2 f(x) dx + \int_2^3 f(x) dx = 1$$

$$\Rightarrow \int_0^1 kx \, dx + \int_1^2 k \, dx + \int_2^3 (-kx + 3k) \, dx = 1$$

$$\left[k \frac{x^2}{2} \right]_0^1 + [kx]_1^2 + \left[-k \cdot \frac{x^2}{2} + 3kx \right]_2^3 = 1 \Rightarrow k = \frac{1}{2}.$$

(ii) For any x such that $-\infty < x < 0$; $F(x) = \int_{-\infty}^x 0 \cdot dt = 0$

$$\text{For any } x, \text{ where } 0 \leq x < 1; F(x) = \int_{-\infty}^0 0 \cdot dt + \int_0^x \frac{t}{2} dt = \frac{x^2}{4}$$

$$\text{For } x, \text{ where } 1 \leq x < 2, F(x) = \int_{-\infty}^0 0 \cdot dt + \int_0^1 \frac{t}{2} dt + \int_1^x \frac{1}{2} dt = \frac{2x-1}{4}$$

For any x , where $2 \leq x < 3$,

$$\begin{aligned} F(x) &= \int_{-\infty}^0 0 \cdot dt + \int_0^1 \frac{t}{2} dt + \int_1^2 \frac{1}{2} dt + \int_2^x \left(-\frac{t}{2} + \frac{3}{2} \right) dt \\ &= \frac{1}{4} + \left(1 - \frac{1}{2} \right) + \left(-\frac{x^2}{4} + \frac{3x}{2} - 2 \right) = -\frac{x^2}{4} + \frac{3x}{2} - \frac{5}{4} \end{aligned}$$

For any x , where $3 \leq x < \infty$

$$\begin{aligned} F(x) &= \int_{-\infty}^0 0 \cdot dt + \int_0^1 \frac{t}{2} dt + \int_1^2 \frac{1}{2} dt + \int_2^3 \left(-\frac{t}{2} + \frac{3}{2} \right) dt + \int_3^x 0 \cdot dt \\ &= \frac{1}{4} + \left(1 - \frac{1}{2} \right) + \left(-\frac{9}{4} + \frac{9}{2} + 1 - 3 \right) = 1 \end{aligned}$$

Hence the distribution function $F(x)$ is given by:

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

(iii) The probability that X is larger than 1.5 is given by:

$$P(X > 1.5) = 1 - P(X < 1.5) = 1 - P(1.5) = \left(1 - \frac{3-1}{4} \right) = \frac{1}{2}.$$

\therefore The probability that X is not larger than 1.5 = $P(X < 1.5) = 1 - \frac{1}{2} = \frac{1}{2}$.

1.5 is; Hence out of three numbers x_1, x_2 and x_3 , the probability that exactly one is larger than

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

Example 7. A petrol pump is supplied with petrol once a day. If its daily volume of sales (X) in thousands of litres is distributed by: $f(x) = 5(1-x)^4$, $0 \leq x \leq 1$, what must be the capacity of its tank in order that the probability that its supply will be exhausted in a given day shall be 0.01?

Sol. Let the capacity of the tank (in '000 of litres) be 'a' such that
 $P(X \geq a) = 0.01$

$$\Rightarrow \int_a^1 f(x) dx = 0.01$$

$$\Rightarrow \int_a^1 5(1-x)^4 dx = 0.01$$

$$\Rightarrow (1-a)^5 = \frac{1}{100}$$

$$\text{or } 1-a = \left(\frac{1}{100}\right)^{1/5}$$

$$\therefore a = 1 - \left(\frac{1}{100}\right)^{1/5} = 1 - 0.3981 = 0.6019$$

Hence the capacity of the tank = $0.6019 \times 1,000$ litres = 601.9 litres.

Example 8. A continuous random variable X has a p.d.f. $f(x) = 3x^2$, $0 \leq x \leq 1$.

Find a and b such that (i) $P(X \leq a) = P(X > a)$, and (ii) $P(X > b) = 0.05$.

Sol. (i) Since $P(X \leq a) = P(X > a)$, each must be equal to $\frac{1}{2}$, because total probability is always unity.

$$\therefore P(X \leq a) = \frac{1}{2}$$

$$\Rightarrow \int_0^a f(x) dx = \frac{1}{2}$$

$$\Rightarrow 3 \int_0^a x^2 dx = \frac{1}{2} \Rightarrow 3 \left| \frac{x^3}{3} \right|_0^a = \frac{1}{2} \Rightarrow a = \left(\frac{1}{2}\right)^{\frac{1}{3}}$$

$$(ii) P(X > b) = 0.05 \Rightarrow \int_b^1 f(x) dx = 0.05$$

$$\Rightarrow 3 \left| \frac{x^3}{3} \right|_b^1 = \frac{1}{20} \Rightarrow 1 - b^3 = \frac{1}{20} \Rightarrow b = \left(\frac{19}{20}\right)^{\frac{1}{3}}.$$

Example 9. (a) A random variable X is distributed at random between the values 0 and 1 so that its probability density function is: $f(x) = kx^2(1-x^3)$, where k is a constant. Find the value of k . Using this value of k , find its mean and variance.

(b) A variable X is distributed at random between the values 0 and 4 and its probability density function is given by: $f(x) = kx^3(4-x)^2$.

Find the value of k , the mean and standard deviation of the distribution.

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

$$k \int_0^1 (x^2 - x^5) dx = 1 \Rightarrow k \left| \frac{x^3}{3} - \frac{x^6}{6} \right|_0^1 = 1 \Rightarrow k = 6.$$

$$\text{Mean} = \mu'_1 = \int_{-\infty}^{\infty} x f(x) dx = 6 \int_0^1 (x^3 - x^6) dx = 6 \left| \frac{x^4}{4} - \frac{x^7}{7} \right|_0^1 = \frac{9}{14}$$

$$\mu'_2 = \int_{-\infty}^{\infty} x^2 f(x) dx = 6 \int_0^1 (x^4 - x^7) dx = 6 \left| \frac{x^5}{5} - \frac{x^8}{8} \right|_0^1 = \frac{9}{20}$$

$$\text{Variance} = \mu_2 = \mu'_2 - \mu'^2_1 = \left\{ \frac{9}{20} - \left(\frac{9}{14} \right)^2 \right\} = \frac{9}{245}.$$

$$(b) \text{ Since } \int_{-\infty}^{\infty} f(x) dx = 1, k \int_0^4 x^3 (4-x)^2 dx = 1 \Rightarrow k = \frac{15}{1024}$$

$$\text{Mean} = \mu'_1 = \int_{-\infty}^{\infty} x f(x) dx = k \int_0^4 x^4 (4-x^2) dx = \frac{16}{7}$$

$$\mu'_2 = \int_{-\infty}^{\infty} x^2 f(x) dx = k \int_0^4 x^5 (4-x)^2 dx = \frac{40}{7}$$

$$\text{Variance} = \mu_2 = \mu'_2 - \mu'^2_1 = \left\{ \frac{40}{7} - \left(\frac{16}{7} \right)^2 \right\} = \frac{24}{49}$$

$$\therefore \text{Standard deviation} = \frac{2\sqrt{6}}{7}.$$

Example 10. The time one has to wait for a bus at a downtown bus stop is observed to be random phenomenon (X) with the following probability density function:

$$f_X(x) = \begin{cases} 0, & \text{for } x < 0 \\ \frac{1}{9}(x+1), & \text{for } 0 \leq x < 1 \\ \frac{4}{9}(x - \frac{1}{2}), & \text{for } 1 \leq x < \frac{3}{2} \\ \frac{4}{9}\left(\frac{5}{2} - x\right), & \text{for } \frac{3}{2} \leq x < 2 \\ \frac{1}{9}(4-x), & \text{for } 2 \leq x < 3 \\ \frac{1}{9}, & \text{for } 3 \leq x < 6 \\ 0, & \text{for } x \geq 6 \end{cases}$$

Let the events A and B be defined as follows:

A: One waits between 0 to 2 minutes inclusive.

B: One waits between 0 to 3 minutes inclusive.

Show that (a) $P(B|A) = \frac{2}{3}$, (b) $P(\bar{A} \cap \bar{B}) = \frac{1}{3}$.

Sol. (a)

$$P(A) = P(X \leq 2) = \int_0^2 f(x) dx$$

$$= \int_0^1 \frac{1}{9}(x+1) dx + \int_1^{3/2} \frac{4}{9} \left(x - \frac{1}{2} \right) dx + \int_{3/2}^2 \frac{4}{9} \left(\frac{5}{2} - x \right) dx$$

$$= \frac{1}{2}. \text{ (on simplification).}$$

$$P(A \cap B) = P[(0 \leq X \leq 2) \cap (1 \leq X \leq 3)] = P(1 \leq X \leq 2) = \int_1^2 f(x) dx$$

$$= \int_1^{3/2} \frac{4}{9} \left(x - \frac{1}{2} \right) dx + \int_{3/2}^2 \frac{4}{9} \left(\frac{5}{2} - x \right) dx$$

$$= \frac{4}{9} \left| \frac{x^2}{2} - \frac{x}{2} \right|_1^{3/2} + \frac{4}{9} \left| \frac{5}{2}x - \frac{x^2}{2} \right|_{3/2}^2 = \frac{1}{3} \text{ (on simplification)}$$

$$\therefore P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}.$$

(b) $\bar{A} \cap \bar{B}$ means that waiting time is more than 3 minutes.

$$\therefore P(\bar{A} \cap \bar{B}) = P(X > 3) = \int_3^{\infty} f(x) dx = \int_3^6 f(x) dx + \int_6^{\infty} f(x) dx \\ = \int_3^6 \frac{1}{9} dx = \frac{1}{9} \left| x \right|_3^6 = \frac{1}{3}$$

TEST YOUR KNOWLEDGE

1. Find the constant k so that function $f(x)$ defined as follows be a density function:

$$f(x) = \begin{cases} 1/k, & a \leq x \leq b \\ 0, & \text{elsewhere} \end{cases}$$

2. Find the value of y_0 so that the function $f(x)$ defined as follows be a density function:

$$f(x) = y_0 e^{-x/\sigma}, \quad 0 \leq x \leq \infty.$$

3. If $f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{otherwise.} \end{cases}$; find the probability $P\left(\frac{1}{4} \leq x \leq \frac{1}{2}\right)$.

4. If $f(x) = \begin{cases} 0, & x < 2 \\ \frac{1}{18}(3 + 2x), & 2 \leq x \leq 4 \\ 0, & x > 4 \end{cases}$

prove that it is a density function. Find the probability that a variate having this density will fall in the interval $2 \leq x \leq 3$.

5. Define distribution function. Verify that the following is a distribution function:

$$F(x) = \begin{cases} 0, & x < -a \\ \frac{1}{2} \left(\frac{x}{a} + 1 \right), & -a \leq x \leq a \\ 1, & x > a. \end{cases}$$

6. What do you understand by probability distribution?

7. Find the mean, variance and the coefficients β_1, β_2 of the distribution: $dF = kx^2 e^{-x} dx = 1, 0 < x < \infty$

8. Show that the symmetrical distribution

$$f(x) = \frac{2a}{\pi} \left(\frac{1}{a^2 + x^2} \right), -a \leq x \leq a$$

represents a probability density function. Also show that

$$\mu_2 = \frac{a^2}{\pi} (4 - \pi) \quad \text{and} \quad \mu_4 = a^4 \left(1 - \frac{8}{3\pi} \right).$$

9. The amount of bread (in hundreds of pounds) x that a certain bakery is able to sell in a day is found to be a numerical valued random phenomenon with a probability function specified by the p.d.f. $f(x)$ is given by

$$f(x) = \begin{cases} kx, & 0 \leq x < 5 \\ k(10 - x), & 5 \leq x < 10 \\ 0, & \text{Otherwise} \end{cases}$$

Find the value of k such that $f(x)$ is a probability density function.

10. The kms X (in thousands of kms) which car owners get with a certain kind of tyre is a random variable having probability density function:

$$f(x) = \begin{cases} \frac{1}{20} e^{-x/20}, & \text{for } x > 0 \\ 0, & \text{for } x \leq 0 \end{cases}$$

Find the probabilities that one of these tyres will last (i) atmost 10,000 kms, (ii) anywhere from 16,000 to 24,000 kms, and (iii) at least 30,000 miles.

[Hint: (i) $P(X \leq 10) = \int_0^{10} f(x) dx$, (ii) $P(16 \leq X \leq 24)$, (iii) $P(X \geq 30)$]

Answers

1. $b - a$
2. $y_0 = \frac{1}{\sigma}$
3. $\frac{3}{16}$
4. $\frac{4}{9}$
7. $k = \frac{1}{2}$, mean = 3, variance = 3, $\beta_1 = \frac{4}{3}$, $\beta_2 = 5$
9. $\frac{1}{25}$
10. (i) 0.3935 (ii) 0.1481 (iii) 0.2231

4.26 JOINT PROBABILITY DISTRIBUTION

Let X and Y be two random variables defined on the same sample space.

Case 1. If X and Y are discrete random variables then a probability function

$$P(X = x_i, Y = y_j) = P(x_i, y_j) \text{ or } p_{ij}$$

that yields the probability that X will assume a particular value x_i while at the same time Y assumes a particular value y_j called a joint probability function of X and Y and has the following properties:

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

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

Case 2. If X and Y are continuous random variables then the probability density function

$$P\left[x - \frac{dx}{2} \leq X \leq x + \frac{dx}{2}, y - \frac{dy}{2} \leq Y \leq y + \frac{dy}{2}\right] = f(x, y)$$

that yields the probability that the point (x, y) lies in the infinitesimal rectangular region of area $dx dy$ is called the joint probability density function of X and Y and satisfies

(i) $f(x, y) \geq 0$ for all (x, y) in the given range.

(ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy$ exists and is equal to 1.

(iii) $P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y) dx dy$

The probability distribution defined in terms of joint probability function or joint probability density function is called a joint probability distribution.

4.27 MARGINAL PROBABILITY DISTRIBUTION

Let X and Y be two discrete random variables defined on the same sample space, such that

$$P(X = x_i) = p_i \text{ and } P(Y = y_j) = p'_j; i = 1, \dots, m, j = 1, \dots, n$$

and

$$P(X = x_i, Y = y_j) = p_{ij}$$

then their joint occurrence have $m \times n$ distinguished pairs (x_i, y_j) which can be arranged in a rectangular array with their probabilities.

As the variable X assumes a definite value x_i , it is accompanied by one of the n values y_1, y_2, \dots, y_n of Y . i.e., $X = x_i$ can occur in n mutually exclusive ways ; therefore

$$P(X = x_i) = P(X = x_i, Y = y_1) + P(X = x_i, Y = y_2) + \dots + P(X = x_i, Y = y_n)$$

or $p = p_{i1} + p_{i2} + \dots + p_{in} = \sum_{j=1}^n p_{ij} = g(x_i)$ say, which is called marginal probability of X for $X = x_i$.

and similarly for fixed y_j

$$p'_j = \sum_{i=1}^n p_{ij} = h(y_j) \text{ say,}$$

which is called marginal probability of Y for $Y = y_j$.

i.e., by adding the probabilities $p_{ij} \dots$ in individual rows and columns, we obtain the marginal probabilities of X and Y which are identical to the individual probabilities of X and Y respectively.

The set of values of a random variable X together with the marginal probabilities is called the *marginal distribution* of that random variable. In fact, marginal distribution is simply the probability distribution of the random variable.

Let X and Y be continuous random variables.

If (X, Y) has the joint density function $f(x, y)$, then the marginal probability density function of X is

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

and marginal probability density function of Y is

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

The corresponding probability distribution is called *Marginal Probability Distribution*.

4.28 CONDITIONAL PROBABILITY DISTRIBUTION

Suppose $P(X = x_i, Y = y_j) = p_{ij}, i = 1, 2, \dots, m, j = 1, 2, \dots, n$ is the joint probability distribution of two discrete random variables, $p_j = P(Y = y_j)$ the probability distribution of Y and $p_i = P(X = x_i)$ the probability distribution of X .

The conditional probability of the event "Y = y_j , given that X = x_i (with $p_i > 0$)" is

$$\frac{P(X = x_i, Y = y_j)}{P(X = x_i)}$$

$$P(Y = y_j | X = x_i) = \frac{p_{ij}}{p_i} = h(y_j/x_i), \text{ say.}$$

The function $h(y_j/x_i) = \frac{p_{ij}}{p_i}$ $i = 1, 2, \dots, m, j = 1, 2, \dots, n$ that yields the conditional probabilities of the random variable Y is called the conditional distribution of Y for given X, and $g(x_i/y_j) = P(X = x_i/Y = y_j) = \frac{p_{ij}}{p_j}$ is the conditional probability distribution of X for given Y.

If (X, Y) has the joint probability density function $f(x, y)$, then the conditional probability density function of X for given Y is defined for any y such that $h(y) > 0$ and is equal to

$$g(x/y) = \frac{f(x, y)}{h(y)}$$

and the conditional p.d.f. of Y for given X is defined for any x such that $g(x) > 0$ and is equal to $h(y/x) = \frac{f(x, y)}{g(x)}$.

ILLUSTRUTIVE EXAMPLES

Example 1. The joint probability distribution of two random variables X and Y is given

by: $P(X = 0, Y = 1) = \frac{1}{3}$, $P(X = 1, Y = -1) = \frac{1}{3}$, and $P(X = 1, Y = 1) = \frac{1}{3}$.

Find (i) Marginal distributions of X and Y, and (ii) the conditional probability distribution of X given $Y = 1$.

$$\begin{aligned} \text{Sol. (i)} \quad P(X = -1) &= \sum_y P(X = -1, Y = y) \\ &= P(X = -1, Y = -1) + P(X = -1, Y = 0) + P(X = -1, Y = 1) = 0 \end{aligned}$$

$$\text{Similarly } P(X = 0) = \frac{1}{3} \quad \text{and} \quad P(X = 1) = \frac{2}{3}$$

X	-1	0	1	Marginal Y
Y				
-1	0	0	$\frac{1}{3}$	$\frac{1}{3}$
0	0	0	0	0
1	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
Marginal (X)	0	$\frac{1}{3}$	$\frac{2}{3}$	1

Thus

Marginal distribution of X is:

Values of X, x :	-1	0	1
$P(X = x)$:	0	$\frac{1}{3}$	$\frac{2}{3}$

Marginal distribution of Y is:

Values of Y, y :	-1	0	1
$P(Y = y)$:	$\frac{1}{3}$	0	$\frac{2}{3}$

(ii) The conditional probability distribution of X given Y is:

$$P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}. \text{ Now}$$

$$P(X = -1 \mid Y = 1) = \frac{P(X = -1, Y = 1)}{P(Y = 1)} = 0,$$

$$P(X = 0 \mid Y = 1) = \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{1/3}{2/3} = \frac{1}{2}$$

$$P(X = 1 \mid Y = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{1/3}{2/3} = \frac{1}{2}$$

Thus the conditional distribution of X given $Y = 1$ is :

Values of $X = x$	-1	0	1
$P(X = x \mid Y = 1)$	0	$\frac{1}{2}$	$\frac{1}{2}$

Example 2. For the adjoining bivariate probability distribution of X and Y, find:

- (i) $P(X \leq 1, Y = 2)$, (ii) $P(X \leq 1)$,
 (iii) $P(Y \leq 3)$, and (iv) $P(X < 3, Y \leq 4)$.

X	Y 1	Y 2	Y 3	Y 4	Y 5	Y 6
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

Sol. The marginal distributions are given below:

X	Y 1	Y 2	Y 3	Y 4	Y 5	Y 6	$p_X(x)$
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{8}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{10}{16}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$	$\frac{8}{64}$
$p_Y(y)$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{11}{64}$	$\frac{13}{64}$	$\frac{6}{32}$	$\frac{16}{64}$	$\Sigma p(x) = 1$ $\Sigma p(y) = 1$

$$(i) P(X \leq 1, Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2) = 0 + \frac{1}{16} = \frac{1}{16}$$

$$(ii) P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{8}{32} + \frac{10}{16} = \frac{7}{8}$$

$$(iii) P(Y \leq 3) = P(Y = 1) + P(Y = 2) + P(Y = 3) = \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$$

$$(iv) P(X < 3, Y \leq 4) = P(X = 0, Y \leq 4) + P(X = 1, Y \leq 4) + P(X = 2, Y \leq 4) \\ = \left(\frac{1}{32} + \frac{2}{32} \right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8} \right) + \left(\frac{1}{32} + \frac{1}{32} + \frac{1}{64} + \frac{1}{64} \right) = \frac{9}{16}.$$

Example 3. For the joint probability distribution of two random variables X and Y given below:

X	Y 1	2	3	4	Total
1	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{10}{36}$
2	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{8}{36}$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	$\frac{9}{36}$
<i>Total</i>	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	1

Find (i) the marginal distributions of X and Y, and (ii) conditional distribution of X given the value of Y = 1 and that of Y given the value of X = 2.

Sol. (i) The marginal distribution of X is defined as:

$$P(X = x) = \sum_y P(X = x, Y = y)$$

$$\therefore P(X = 1) = \sum_y P(X = 1, Y = y) \\ = P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 1, Y = 4) \\ = \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36}.$$

$$\text{Similarly, } P(X = 2) = \sum_y P(X = 2, Y = y) = \frac{9}{36}; P(X = 3) = \sum_y P(X = 3, Y = y) = \frac{8}{36}$$

and

$$P(X = 4) = \sum_y P(X = 4, Y = y) = \frac{9}{36}.$$

Similarly, we can obtain the marginal distribution of Y.

Marginal Distribution of X

Values of X, x	1	2	3	4
Values of Y, y	1	2	3	4
P(X = x)	$\frac{10}{36}$	$\frac{9}{36}$	$\frac{8}{36}$	$\frac{9}{36}$

Marginal Distribution of Y

Values of Y, y	1	2	3	4
Values of X, x	1	2	3	4
P(Y = y)	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$

(ii) The conditional probability function of X given Y is defined as follows:

$$P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}. \text{ Therefore}$$

$$P(X = 1 \mid Y = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{4/36}{11/36} = \frac{4}{11}$$

$$P(X = 2 \mid Y = 1) = \frac{P(X = 2, Y = 1)}{P(Y = 1)} = \frac{1/36}{11/36} = \frac{1}{11}$$

$$P(X = 3 \mid Y = 1) = \frac{P(X = 3, Y = 1)}{P(Y = 1)} = \frac{5/36}{11/36} = \frac{5}{11}$$

$$P(X = 4 \mid Y = 1) = \frac{P(X = 4, Y = 1)}{P(Y = 1)} = \frac{1/36}{11/36} = \frac{1}{11}$$

Hence the conditional distribution of X given Y = 1 is:

$$\begin{array}{cccc} x : & 1 & 2 & 3 & 4 \\ P(X = x \mid Y = 1) : & \frac{4}{11} & \frac{1}{11} & \frac{5}{11} & \frac{1}{11} \end{array}$$

Similarly, we can obtain the conditional distribution of Y for X = 2 as given below:

$$\begin{array}{cccc} y : & 1 & 2 & 3 & 4 \\ P(Y = y \mid X = 2) : & \frac{1}{9} & \frac{1}{3} & \frac{1}{3} & \frac{2}{9} \end{array}$$

Example 4. Two discrete random variables X and Y have the joint probability density function:

$$p_{XY}(x, y) = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!}, \quad y = 0, 1, 2, \dots, x; x = 0, 1, 2, \dots,$$

where λ, p are constants with $\lambda > 0$ and $0 < p < 1$.

Find (i) The marginal probability density functions of X and Y.

(ii) The conditional distribution of Y for a given X and of X for a given Y.

Sol. (i) The marginal p.m.f. of X is given by:

$$\begin{aligned} p_X(x) &= \sum_{y=0}^x p(x, y) = \sum_{y=0}^x \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} = \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^x \frac{x! p^y (1-p)^{x-y}}{y!(x-y)!} \\ &= \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^x {}^x C_y p^y (1-p)^{x-y} = \frac{\lambda^x e^{-\lambda}}{x!} [p + (1-p)]^x = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots \end{aligned}$$

which is the probability function of a Poisson distribution with parameter λ .

$$\begin{aligned} p_Y(y) &= \sum_{x=y}^{\infty} p(x, y) = \sum_{x=y}^{\infty} \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} = \frac{(\lambda p)^y e^{-\lambda}}{y!} \sum_{x=y}^{\infty} \frac{[\lambda(1-p)]^{x-y}}{(x-y)!} \\ &\quad [\because y = 0, 1, 2, \dots, x \Rightarrow x \leq y \Rightarrow x \geq y] \\ &= \frac{(\lambda p)^y e^{-\lambda}}{y!} e^{\lambda(1-p)} = \frac{e^{-\lambda p} (\lambda p)^y}{y!}; \quad y = 0, 1, 2, \dots \end{aligned}$$

which is the probability function of a Poisson distribution with parameter λp .

(ii) The conditional distribution of Y for given X is:

$$\begin{aligned} P_{Y|X}(y|x) &= \frac{p_{XY}(x,y)}{p_X(x)} = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y} x!}{y!(x-y)! \lambda^x e^{-\lambda}} = \frac{x!}{y!(x-y)!} p^y (1-p)^{x-y} \\ &= {}^x C_y p^y (1-p)^{x-y}, \quad x \geq y \quad i.e., \quad y = 0, 1, 2, \dots, x. \end{aligned}$$

The conditional probability distribution of X for given Y is:

$$\begin{aligned} p_{X|Y}(x|y) &= \frac{p_{XY}(x,y)}{p_Y(y)} = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} \cdot \frac{y!}{e^{-\lambda p} (\lambda p)^y} \quad [c.f. \text{ Part (i)}] \\ &= \frac{e^{-\lambda q} (\lambda q)^{x-y}}{(x-y)!}; \quad q = 1-p, \quad x \geq y \quad i.e., \quad x = y, y+1, y+2, \dots \end{aligned}$$

Example 5. If X and Y are two random variables having joint density function:

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

Find (i) $P(X < 1 \cap Y < 3)$, (ii) $P(X + Y < 3)$, and (iii) $P(X < 1 \mid Y < 3)$.

Sol. We have

$$(i) P(X < 1 \cap Y < 3) = \int_{-\infty}^1 \int_{-\infty}^3 f(x,y) dx dy = \int_0^1 \int_2^3 \frac{1}{8}(6-x-y) dx dy = \frac{3}{8}$$

$$(ii) P(X + Y < 3) = \int_0^1 \int_2^{3-x} \frac{1}{8}(6-x-y) dx dy = \frac{5}{24}$$

$$(iii) P(X < 1 \mid Y < 3) = \frac{P(X < 1 \cap Y < 3)}{P(Y < 3)} = \frac{3/8}{5/8} = \frac{3}{5}$$

$$\left[\text{From part (i) and } P(Y < 3) = \int_0^2 \int_2^3 \frac{1}{8}(6-x-y) dx dy = \frac{5}{8} \right]$$

Example 6. Suppose that two-dimensional continuous random variable (X, Y) has joint

$$\text{p.d.f. given by } f(x,y) = \begin{cases} 6x^2 y, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$(i) \text{ Verify that } \int_0^1 \int_0^1 f(x,y) dx dy = 1.$$

$$(ii) \text{ Find } P\left(0 < X < \frac{3}{4}, \frac{1}{3} < Y < 2\right), P(X + Y < 1), P(X > Y) \text{ and } P(X < 1 \mid Y < 2).$$

$$\text{Sol. (i)} \int_0^1 \int_0^1 f(x,y) dx dy = \int_0^1 \int_0^1 6x^2 y dx dy = \int_0^1 6x^2 \left| \frac{y^2}{2} \right|_0^1 dx = \int_0^1 3x^2 dx = \left| x^3 \right|_0^1 = 1$$

$$\begin{aligned} (ii) \quad P(0 < X < \frac{3}{4}, \frac{1}{3} < Y < 2) &= \int_0^{3/4} \int_{1/3}^1 6x^2 y dx dy + \int_0^{3/4} \int_1^2 0 dx dy \\ &= \int_0^{3/4} 6x^2 \left| \frac{y^2}{2} \right|_{1/3}^1 dx = \frac{8}{9} \int_0^{3/4} 3x^2 dx = \frac{8}{9} \left| x^3 \right|_0^{3/4} = \frac{3}{8}. \end{aligned}$$

$$P(X + Y < 1) = \int_0^1 \int_1^{1-x} 6x^2 y \, dx dy = \int_0^1 6x^2 \left| \frac{y^2}{2} \right|_0^{1-x} dx = \int_0^1 3x^2(1-x)^2 dx = \frac{1}{10}$$

$$P(X > Y) = \int_0^1 \int_0^x 6x^2 y \, dx dy = \int_0^1 3x^2 \left| y^2 \right|_0^x dx = \int_0^1 3x^4 dx = \frac{3}{5}.$$

$$P(X < 1 \mid Y < 2) = \frac{P(X < 1 \cap Y < 2)}{P(Y < 2)}$$

where $P(X < 1 \cap Y < 2) = \int_0^1 \int_0^1 6x^2 y \, dx dy + \int_0^1 \int_1^2 0 \cdot dx dy = 1$

and $P(Y < 2) = \int_0^1 \int_0^2 f(x, y) \, dx dy = \int_0^1 \int_0^1 6x^2 y \, dx dy + \int_0^1 \int_1^2 0 \cdot dx dy = 1$

$$\therefore P(X < 1 \mid Y < 2) = \frac{P(X < 1 \cap Y < 2)}{P(Y < 2)} = 1.$$

Example 7. The joint probability density function of a two-dimensional random variable (X, Y) is given by:

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

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

(ii) Find the conditional density function of Y given $X = x$ and conditional density function of X given $Y = y$.

(iii) Check for independence of X and Y .

Sol. Evidently $f(x, y) \geq 0$ and $\int_0^1 \int_0^x 2 \, dx dy = 2 \int_0^1 x \, dx = 1$.

(i) The marginal p.d.f.'s of X and Y are given by

$$f_X(x) = \begin{cases} \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy = \int_0^x 2 \, dy = 2x, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx = \int_y^1 2 \, dx = 2(1-y), & 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

(ii) The conditional density function of Y given X , ($0 < x < 1$) is

$$f_{Y \mid X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x}, \quad 0 < y < x.$$

The conditional density function of X given Y , ($0 < y < 1$) is

$$f_{X \mid Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{2}{2(1-y)} = \frac{1}{(1-y)}, \quad y < x < 1$$

(iii) Since $f_X(x)f_Y(y) = 2(2x)(1-y) \neq f_{XY}(x, y)$, X and Y are not independent.

Example 8. The joint p.d.f. of two random variables X and Y is given by

$$f(x, y) = \frac{9(1+x+y)}{2(1+x)^4(1+y)^4}; \quad 0 \leq x < \infty, \quad 0 \leq y < \infty$$

Find the marginal distributions of X and Y , and the conditional distribution of Y for $X=x$.

Sol. Marginal p.d.f. of X is given by:

$$\begin{aligned} f_X(x) &= \int_0^\infty f(x, y) dy = \frac{9}{2(1+x)^4} \int_0^\infty \frac{(1+y)+x}{(1+y)^4} dy \\ &= \frac{9}{2(1+x)^4} \int_0^\infty \left\{ (1+y)^{-3} + x(1+y)^{-4} \right\} dy \\ &= \frac{9}{2(1+x)^4} \left(\left| \frac{-1}{2(1+y)^2} \right|_0^\infty + x \left| \frac{-1}{3(1+y)^3} \right|_0^\infty \right) \\ &= \frac{9}{2(1+x)^4} \cdot \left(\frac{1}{2} + \frac{x}{3} \right) = \frac{3}{4} \cdot \frac{3+2x}{(1+x)^4}; 0 < x < \infty \end{aligned}$$

Since $f(x, y)$ is symmetric in x and y, the marginal p.d.f. of Y is given by:

$$f_Y(y) = \int_0^\infty f(x, y) dx = \frac{3}{4} \cdot \frac{3+2y}{(1+y)^4}; 0 < y < \infty$$

The conditional distribution of Y for $X = x$ is given by:

$$f_{XY}(Y=y | X=x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{9(1+x+y)}{2(1+x)^4(1+y)^4} \times \frac{4(1+x)^4}{3(3+2x)} = \frac{6(1+x+y)}{(1+y)^4(3+2x)}; 0 < y < \infty.$$

Example 9. Joint distribution of X and Y is given by: $f(x, y) = 4xy e^{-(x^2 + y^2)}$; $x \geq 0, y \geq 0$.

Test whether X and Y are independent. For the above joint distribution, find the conditional density of X given $Y = y$.

Sol. Marginal density of X is given by

$$\begin{aligned} f_X(x) &= \int_0^\infty f_{XY}(x, y) dy = \int_0^\infty 4xy e^{-(x^2 + y^2)} dy = 4x e^{-x^2} \int_0^\infty y e^{-y^2} dy \\ &= 4x e^{-x^2} \cdot \int_0^\infty e^{-t} \cdot \frac{dt}{2} = 2x \cdot e^{-x^2} \Big| -e^{-t} \Big|_0^\infty \\ \therefore f_X(x) &= 2x e^{-x^2}; x \geq 0 \end{aligned}$$

Similarly, the marginal p.d.f. of Y is given by:

$$f_Y(y) = \int_0^\infty f_{XY}(x, y) dx = 2y e^{-y^2}; y \geq 0$$

Since $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$, X and Y are independently distributed. The conditional distribution of X for given Y is given by :

$$f_{X|Y}(X=x | Y=y) = \frac{f(x, y)}{f_Y(y)} = 2x e^{-x^2}; x \geq 0.$$

Example 10. Let X and Y be jointly distributed with p.d.f.:

$$f_{XY}(x, y) = \begin{cases} \frac{1}{4} (1+xy), & |x| < 1, |y| < 1 \\ 0, & \text{otherwise} \end{cases}$$

Show that X and Y are not independent but X^2 and Y^2 are independent.

Sol. $f_X(x) = \int_{-1}^1 f(x, y) dy = \frac{1}{4} \left| y + \frac{xy^2}{2} \right|_{-1}^1 = \frac{1}{2}, -1 < x < 1;$

Similarly, $f_Y(y) = \int_{-1}^1 f(x, y) dx = \frac{1}{2}, -1 < y < 1$

Since $f_{X, Y}(x, y) \neq f_X(x) f_Y(y)$, X and Y are not independent. However,

$$P(X^2 \leq x) = P(|X| \leq \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_X(x) dx = \sqrt{x} \quad \dots(1)$$

$$\begin{aligned} P(X^2 \leq x \cap Y^2 \leq y) &= P(|X| \leq \sqrt{x} \cap |Y| \leq \sqrt{y}) \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} \left[\int_{-\sqrt{y}}^{\sqrt{y}} f(u, v) dv \right] du = \sqrt{x} \sqrt{y} \\ &= P(X^2 \leq x) \cdot P(Y^2 \leq y) \end{aligned}$$

[From (1)]

Hence, X^2 and Y^2 are independent.

TEST YOUR KNOWLEDGE

1. Find the mean and variance of p.d.f.

$$f(x) = \begin{cases} \frac{1}{4} e^{-x/4}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

2. Find the mean and variance of exponential distribution $f(x) = \frac{1}{b} e^{-x/b}; x > 0, b > 0$

3. A continuous random variable X has the probability density function:

$$f(x) = A + Bx, 0 \leq x \leq 1.$$

If the mean of the distribution is $\frac{1}{2}$, find A and B.

4. For the probability density function: $f(x) = cx^2(1-x)$, $0 < x < 1$, find
(i) the constant c, and (ii) mean.

5. The distribution of a continuous random variable X in range $(-3, 3)$ is given by p.d.f.:

$$f(x) = \begin{cases} \frac{1}{16} (3+x)^2, & -3 \leq x \leq -1 \\ \frac{1}{16} (6-2x^2), & -1 \leq x \leq 1 \\ \frac{1}{16} (3-x)^2, & 1 \leq x \leq 3 \end{cases}$$

(i) Verify that the area under the curve is unity.

(ii) Find the mean and variance of the above distribution.

6. The length of time (in minutes) that a certain lady speaks on the telephone is found to be random phenomenon, with a probability function specified by the probability density function $f(x)$ as:

$$f(x) = \begin{cases} A e^{-x/5}, & \text{for } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the value of A that makes $f(x)$ a p.d.f.

(b) What is the probability that the number of minutes that she will talk over the phone is:
(i) more than 10 minutes, (ii) less than 5 minutes, and (iii) between 5 and 10 minutes.

1. (a) Find k so that $f(x, y) = kxy$, $1 \leq x \leq y \leq 2$ will be a probability density function.

(b) If $f(x, y) = \begin{cases} e^{-(x+y)}, & x \geq 0, y \geq 0 \\ 0, & \text{elsewhere} \end{cases}$ is the joint probability density function of random variables X and Y , find

(i) $P(X < 1)$, (ii) $P(X > Y)$, and (iii) $P(X + Y < 1)$.

8. The joint probability density function of the two-dimensional variable (X, Y) is of the form:

$$f(x, y) = \begin{cases} e^{-(x+y)}, & 0 \leq y < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

(i) Determine the constant k .

(ii) Find the conditional probability density function $f_1(x | y)$.

(iii) Compute $P(Y \geq 3)$.

9. Two-dimensional random variable (X, Y) have the joint density function

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

Find the marginal and conditional distributions.

10. (a) The joint probability density function of the two-dimensional random variable (X, Y) is given by:

$$f(x, y) = \begin{cases} \frac{8}{9}xy, & 1 \leq x \leq y \leq 2 \\ 0, & \text{elsewhere} \end{cases}$$

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

(ii) Find the conditional density function of Y given $X = x$, and conditional density function of X given $Y = y$.

(b) Let $f(x_1, x_2) = \begin{cases} A(x_1 x_2 + e^{x_1}); & 0 < (x_1 x_2) < 1 \\ 0, & \text{elsewhere} \end{cases}$ be the joint p.d.f. of (X_1, X_2) . Determine A .

11. The joint probability density function of the two dimensional random variable (X, Y) is given by :

$$f(x, y) = \begin{cases} x^3 y^3 / 16, & 0 \leq x \leq 2, 0 \leq y \leq 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find the marginal densities of X and Y . Also, find the cumulative distribution functions for X and Y .

12. (a) If the joint distribution function of X and Y is given by

$$F_{XY}(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y)}; & x > 0, y > 0 \\ 0; & \text{elsewhere} \end{cases}$$

(i) Find the marginal densities of X and Y .

(ii) Are X and Y independent?

(iii) Find $P(X \leq 1 \cap Y \leq 1)$ and $P(X + Y \leq 1)$.

(b) Given $f(x, y) = e^{-(x+y)}$, $0 \leq x < \infty$, $0 \leq y < \infty$. Are X and Y independent? Also, find $P(X > 1)$.

13. A two-dimensional random variable (X, Y) have a bivariate distribution given by :

$$P(X=x, Y=y) = \frac{x^2 + y}{32}, \text{ for } x = 0, 1, 2, 3 \text{ and } y = 0, 1. \text{ Find the marginal distributions of } X \text{ and } Y.$$

14. A two-dimensional random variable (X, Y) have a joint probability mass function:

$$p(x, y) = \frac{1}{27} (2x + y), \text{ where } x \text{ and } y \text{ can assume only the integer values 0, 1 and 2.}$$

Find the conditional distribution of Y for X = x.

Answers

1. Mean = 4, Variance = 80

2. Mean = b, Variance = b²

3. A = 1, B = 0

4. (i) c = 12, (ii) Mean = 3/5

5. (ii) Mean = 0, Variance = 1

6. (a) A = $\frac{1}{5}$ (b) (i) $\frac{1}{e^2}$ (ii) $\frac{e-1}{e}$ (iii) $\frac{e-1}{e^2}$

7. (a) k = 8/9 (b) (i) $1 - \frac{1}{e}$ (ii) $\frac{1}{2}$ (iii) $1 - \frac{2}{e}$

8. (i) k = 1 (ii) $f_1(x|y) = e^{-x}$ (iii) e^{-3}

9. (i) $f_X(x) = \begin{cases} 4x(1-x^2), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}, f_Y(y) = 4y^3, 0 < y < 1$

(ii) $f_{X|Y}(x|y) = \frac{2x}{y^2}, 0 < x < y, 0 < y < 1$

$f_{Y|X}(y|x) = \frac{2y}{1-x^2}, x < y < 1, 0 < x < 1$

10. (a) (i) $f_X(x) = \begin{cases} \frac{4}{9}x(4-x^2), & 1 \leq x \leq 2 \\ 0, & \text{otherwise} \end{cases}$

(ii) $f_{X|Y}(x|y) = \frac{2x}{y^2-1}, 1 \leq x \leq y$

$f_Y(y) = \frac{4}{9}y(y^2-1), 1 \leq y \leq 2$

$f_{Y|X}(y|x) = \frac{2y}{4-x^2}, x \leq y \leq 2.$

(b) A = $\frac{4}{4e-3}$.

11. $f_X(x) = \frac{x^3}{4}, 0 \leq x \leq 2, f_Y(y) = \frac{y^3}{4}, 0 \leq y \leq 2; F_X(x) = \begin{cases} 0, x < 0 \\ x^4/16, 0 \leq x \leq 2, \\ 1, x > 2 \end{cases} F_Y(y) = \begin{cases} 0, y < 0 \\ y^4/16, 0 \leq y \leq 2 \\ 1, y > 2 \end{cases}$

12. (a) (i) $f_X(x) = e^{-x}, x \geq 0; f_Y(y) = e^{-y}, y \geq 0$ (ii) Yes

(iii) $P(X \leq 1 \cap Y \leq 1) = \left(1 - \frac{1}{e}\right)^2, P(X + Y \leq 1) = 1 - \frac{2}{e}$

(b) Yes; $\frac{1}{e}$.

X Y	0	1	2	3	Marginal distribution of Y, P(Y = y)
0	0	1/32	1/8	9/32	7/16
1	1/32	1/16	5/32	5/16	9/16
Marginal distribution of X, P(X = x)	1/32	3/32	9/32	19/32	1

14. Conditional distribution of Y for X = x.

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

4.29 MATHEMATICAL EXPECTATION OR EXPECTED VALUE OF A RANDOM VARIABLE

Once we have constructed the probability distribution for a random variable, we often want to compute the mean or expected value of the random variable. The *expected value* of a discrete random variable is a weighted average of all possible values of the random variable, where the weights are the probabilities associated with the corresponding values.

If x , denotes a discrete random variable which can assume the values x_1, x_2, \dots, x_n with respective probabilities p_1, p_2, \dots, p_n where $p_1 + p_2 + \dots + p_n = 1$, the *mathematical expectation* of x or simply the *expectation* of X , denoted by $E(x)$, is defined as

$$E(x) = p_1x_1 + p_2x_2 + \dots + p_nx_n = \sum_{i=1}^n p_i x_i = \Sigma p x, \Sigma p = 1.$$

Let $\phi(x)$ be a function of the variate x so that it takes the values $\phi(x_1), \phi(x_2), \dots$. When x takes the values x_1, x_2, \dots ; and if p_1, p_2, \dots , be the respective probabilities, then the expected value of the function $\phi(x)$ is defined as

$$E[\phi(x)] = p_1\phi(x_1) + p_2\phi(x_2) + \dots + p_n\phi(x_n), \Sigma p = 1$$

If $\phi(x) = x^r$, then

$$E(x^r) = p_1x_1^r + p_2x_2^r + \dots + p_nx_n^r.$$

This is defined as the r^{th} moment of the discrete probability distribution about $x = 0$ and is denoted by $\mu r'$. This $\mu r'$ is the expected value of the r^{th} power of the variate.

Also

$$\mu r = E(x - \bar{x})^r = \Sigma p_i(x_i - \bar{x})^r$$

In particular

$$\mu_1' = E(x) = p_1x_1 + p_2x_2 + \dots = \sum_i p_i x_i$$

Here if p_i is replaced by $\frac{f_i}{N}$ where $\Sigma f_i = N$, then

$$E(x) = \frac{\Sigma f x}{N}, \text{ which is the mean.}$$

$\therefore E(x)$ represents the mean.

Just as in the case of frequency, this moment is called the *mean value* of the variate or of the distributions. More generally it is known as the *expected value* or the *expectation* of the variate x . The relation

$$\mu_2 = E[(x - E(x))^2] = E(x^2) - [E(x)]^2.$$

is called the *variance* of the distribution of x and is denoted by $\text{var}(x)$. The relation between μ_r and μ_r' are the same as for a frequency distribution. But here we have

$$E[x - E(x)] = 0$$

Thus the expected value of the deviation of variate from its mean is zero.

If x is a continuous random variable with probability density function $f(x)$, then

$$E(x) = \int_{-\infty}^{\infty} xf(x) dx = \int_{-\infty}^{\infty} x dF(x)$$

is expectation of continuous random variable with certain restrictions.

4.30 LAWS OF EXPECTATION

Theorem 1. If C is a constant then $E(C) = C$.

Proof. Since the constant function C , assigns the value C to each value within its domain, we have

$$P(C = C) = 1$$

$$P(C = D) = 0, D \neq C$$

therefore

$$E(C) = C \cdot 1 + D \times 0 = C.$$

Theorem 2. If a is a constant, then $E(aX) = aE(X)$.

Proof. Let X take values x_1, x_2, \dots with probabilities p_1, p_2, \dots

$$\text{Then } E(aX) = ax_1p_1 + ax_2p_2 + \dots \quad \{ \because P(aX = ax_i) = P(X = x_i)\}$$

$$= a(x_1p_1 + x_2p_2 + \dots) = aE(X).$$

Theorem 3. Expectation of the sum of two Random Variables.

If X and Y are two discrete random variables with finite expectations $E(X)$ and $E(Y)$ respectively, then the expectation of their sum exists and is the sum of their expectations i.e.,

$$E(X + Y) = E(X) + E(Y).$$

Corollary 1. If X_1, X_2, \dots, X_n are n random variables with finite expectations $E(X_1), E(X_2), \dots, E(X_n)$, then the expectation of their sum exists and is the sum of their expectations. In symbols

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

or

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

i.e., the sign E and Σ are interchangeable.

Corollary 2. If a and b are constants, then $E(aX + b) = aE(X) + b$.

$$\begin{aligned} \text{Proof. } E(aX + b) &= E(aX) + E(b), & \{ \because E(X + Y) = E(X) + E(Y)\} \\ &= aE(X) + b. & \{ \because E(aX) = aE(X)\} \end{aligned}$$

Theorem 4. Expectation of the Product of two Independent Random Variables.

If X and Y are two independent discrete random variables with finite expectations, then the expectation of their product exists and is equal to the product of their expectations. In symbols,

$$E(XY) = E(X) E(Y).$$

Corollary. If X_1, X_2, \dots, X_k are mutually independent random variables with finite expectations, then the expectation of their product exists, and is equal to the product of their expectations. In symbols

$$E(X_1, X_2, \dots, X_k) = E(X_1) E(X_2) \dots E(X_k).$$

Theorem 5. *Expectation of the difference of two random variables.*
 If X and Y are two random variables with finite expectations $E(X)$ and $E(Y)$ respectively then the expectation of their difference is equal to the difference of their expectations. In symbols

$$E(X - Y) = E(X) - E(Y).$$

In general

$$E(X_1 \pm X_2 \pm \dots \pm X_k) = E(X_1) \pm E(X_2) \pm \dots \pm E(X_k).$$

ILLUSTRATIVE EXAMPLES

Example 1. What is the expected value of the number of points that will be obtained in a single throw with an ordinary die? Find variance also.

Sol. Here the variate is the number of points showing on a die. It assumes the values 1, 2, 3, 4, 5, 6 with probability $\frac{1}{6}$ in each case.

Hence,

$$E(x) = p_1x_1 + p_2x_2 + \dots + p_6x_6 = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \dots + \frac{1}{6} \cdot 6 = 3.5$$

Also,

$$\text{var}(x) = E(x^2) - [E(x)]^2 = \frac{1}{6}(1^2 + 2^2 + \dots + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Example 2. Thirteen cards are drawn simultaneously from a deck of 52. If aces count 1, face cards 10 and others according to denomination, find the expectation of the total score on the 13 cards.

Sol. Let x_i be the number corresponding to the i^{th} card, then x_i takes the values 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 10, 10; 10 each having the probability $\frac{1}{13}$. Hence

$$E(x_i) = \frac{1}{13} \cdot 1 + \frac{1}{13} \cdot 2 + \frac{1}{13} \cdot 3 + \dots + \frac{1}{13} \cdot 9 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10$$

$$E(x_i) = \frac{1}{13} (1 + 2 + 3 + \dots + 9 + 10 + 10 + 10 + 10) = \frac{85}{13}.$$

Example 3. In four tosses of a coin, let x be the number of heads. Calculate the expected values of x .

Sol. There will be all the heads, there is only one way

i.e.,

$$P(x = 4) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$$

$$P(x = 3) = \frac{^4C_3}{16} = \frac{4}{16} = \frac{1}{4}$$

$$P(x = 2) = \frac{^4C_2}{16} = \frac{6}{16} = \frac{3}{8}$$

$$P(x = 1) = \frac{^4C_1}{16} = \frac{4}{16} = \frac{1}{4}$$

$$P(x = 0) = \frac{^4C_0}{16} = \frac{1}{16}.$$

Hence the probability distribution of x is

$x:$	0	1	2	3	4
$P(x):$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

$$\therefore E(x) = \sum_{x=0}^4 x(p_x) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16} = 2.$$

Example 4. Find $E(x)$, $E(x^2)$, $E\{(x - \bar{x})^2\}$ for the following probability distribution:

$x:$	8	12	16	20	24
$p(x):$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{12}$

$$E(x) = \sum x \cdot p(x) = 8 \cdot \frac{1}{8} + 12 \cdot \frac{1}{6} + 16 \cdot \frac{3}{8} + 20 \cdot \frac{1}{4} + 24 \cdot \frac{1}{12} = 16.$$

This represents the *mean* of the distribution.

$$E(x^2) = \sum x^2 \cdot p(x) = 8^2 \cdot \frac{1}{8} + 12^2 \cdot \frac{1}{6} + 16^2 \cdot \frac{3}{8} + 20^2 \cdot \frac{1}{4} + 24^2 \cdot \frac{1}{12} = 276.$$

This represents the *second moment* about the origin zero.

$$\begin{aligned} E\{(x - \bar{x})^2\} &= \sum (x - \bar{x})^2 p(x) \\ &= (8 - 16)^2 \cdot \frac{1}{8} + (12 - 16)^2 \cdot \frac{1}{6} + (16 - 16)^2 \cdot \frac{3}{8} + (20 - 16)^2 \cdot \frac{1}{4} + (24 - 16)^2 \cdot \frac{1}{12} = 20 \end{aligned}$$

This represents the *variance of the distribution*.

4.31 INDEPENDENT VARIATES

The variates, x and y , are said to be independent, if the probability that either of them will assume a prescribed value does not depend on the value assumed by the other.

3.31.1 Theorem. The expectation of the product of two independent variates is equal to the product of their expectations i.e., $E(xy) = E(x)E(y)$.

4.32 COVARIANCE

If x and y are two variates with the respective expected values (or means) x and y , the covariance between x and y is defined as

$$\text{cov}(x, y) = E[(x - \bar{x})(y - \bar{y})].$$

Thus the expected value of product of the deviations of the two variates from their mean is called their *covariance*.

4.33 THEOREM

The covariance of two independent variates is equal to zero.

If x and y are independent variates, then

$$E(x - \bar{x}) = E(x) - E(\bar{x}) = \bar{x} - \bar{x} = 0,$$

$$E(y - \bar{y}) = E(y) - E(\bar{y}) = \bar{y} - \bar{y} = 0.$$

$$\therefore \text{cov}(x, y) = E(x - \bar{x})E(y - \bar{y}) = 0$$

Note.

$$\begin{aligned}\text{cov}(x, y) &= E\{(x - \bar{x})(y - \bar{y})\} = E(xy - \bar{x}y - \bar{y}x + \bar{x}\bar{y}) \\&= E(xy) - E(\bar{x}y) - E(\bar{y}x) + E(\bar{x}\bar{y}) \\&= E(xy) - \bar{x}E(y) - \bar{y}E(x) + \bar{x}\bar{y} \\&= E(xy) - \bar{x}\bar{y} - \bar{y}\bar{x} + \bar{x}\bar{y} \\&= E(xy) - E(x)E(y).\end{aligned}$$

1.34 VARIANCE OF n VARIATES

x_1, x_2, \dots, x_n be n variates with variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ to find the variance of u where

$$u = a_1x_1 + a_2x_2 + \dots + a_nx_n \text{ where } a's \text{ are constant.}$$

$$\begin{aligned}E(u) &= E(a_1x_1 + a_2x_2 + \dots + a_nx_n) \\&= E(a_1x_1) + E(a_2x_2) + \dots + E(a_nx_n) \\&= a_1E(x_1) + a_2E(x_2) + \dots + a_nE(x_n)\end{aligned}$$

$$\therefore u - E(u) = a_1[x_1 - E(x_1)] + a_2[x_2 - E(x_2)] + \dots + a_n[x_n - E(x_n)]$$

Squaring both sides, we get

$$\begin{aligned}[u - E(u)]^2 &= a_1^2[x_1 - E(x_1)]^2 + a_2^2[x_2 - E(x_2)]^2 + \dots \\&\quad + a_n^2[x_n - E(x_n)]^2 + 2a_1a_2[x_1 - E(x_1)][x_2 - E(x_2)] + \dots\end{aligned}$$

Taking expected values of both sides, we get

$$\begin{aligned}\text{var}(u) &= a_1^2 \text{var}(x_1) + a_2^2 \text{var}(x_2) + \dots + a_n^2 \text{var}(x_n) \\&\quad + 2a_1a_2 \text{cov}(x_1, x_2) + \dots + 2a_{n-1}a_n \text{cov}(x_{n-1}, x_n)\end{aligned}$$

Cor. 1. If $a_1 = a_2 = 1, a_3 = a_4 = \dots = a_n = 0$, we have

$$\text{var}(x_1 + x_2) = \text{var}(x_1) + \text{var}(x_2) + 2 \text{cov}(x_1, x_2).$$

Cor. 2. If $a_1 = 1, a_2 = -1, a_3 = a_4 = \dots = a_n = 0$, we have

$$\text{var}(x_1 - x_2) = \text{var}(x_1) + \text{var}(x_2) - 2 \text{cov}(x_1, x_2).$$

Cor. 3. If x_1, x_2, \dots, x_n are independent, we have

$$\text{var}(u) = a_1^2 \text{var}(x_1) + a_2^2 \text{var}(x_2) + \dots + a_n^2 \text{var}(x_n).$$

ILLUSTRATIVE EXAMPLES

Example 1. Balls are taken one by one out of an urn containing a white and b black balls until the first white ball is drawn. Show that the expectation of the number of black balls

preceding the first white ball is $\frac{b}{a+1}$.

Sol. Let x_1 denote the number of black balls before the first white ball is drawn and x_i denote the number of black balls following the $(i-1)^{\text{th}}$ white balls and preceding the i^{th} white ball.

Let x_{a+1} denote the number of black balls left after the white balls have been drawn.

Then

Hence

or

$$x_1 + x_2 + \dots + x_{a+1} = b,$$

$$E(x_1 + x_2 + \dots + x_{a+1}) = E(b),$$

$$E(x_1) + E(x_2) + \dots + E(x_{a+1}) = E(b).$$

The probability of every system of numbers x_1, x_2, \dots, x_{a+1} is the same and is $\frac{a!b!}{(a+b)!}$

Thus, $E(x_1) = E(x_2) = \dots = E(x_{a+1}) = K$ say.

Therefore, $(a+1)K = b$ giving $K = \frac{b}{a+1}$.

Example 2. A box contains a white and b black balls; c balls are drawn. Show that the expectation of the number of white balls drawn is $\frac{ca}{a+b}$.

Sol. Let a variate x_i be defined as follows

$$x_i = x_1, x_2, x_3, \dots, x_c$$

and in that case

$$x_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ ball drawn is white} \\ 0 & \text{if the } i^{\text{th}} \text{ ball drawn is black} \end{cases}$$

Then the number of white balls, s , among the c drawn balls is given by

$$s = x_1 + x_2 + \dots + x_c.$$

$$\text{Now, } P(x_i = 1) = \frac{a}{a+b}, P(x_i = 0) = \frac{b}{a+b}$$

$$\therefore E_i = 1 \cdot \frac{a}{a+b} + 0 \cdot \frac{b}{a+b} = \frac{a}{a+b}$$

$$\therefore E(S) = \frac{a}{a+b} + \frac{a}{a+b} + \dots + \frac{a}{a+b}, c \text{ terms} = \frac{ca}{a+b}$$

Example 3. What is the expectation of the number of failures preceding the first success in an infinite series of independent trials with constant probability p of success in a trial?

Sol. Let x denote the number of failures preceding the first success we find that x takes the values $0, 1, 2, 3, 4, \dots$. Their respective probabilities are $p, (1-p)p, (1-p)^2p, (1-p)^3p, \dots$

Since, $1-p = q$, these are p, qp, q^2p, q^3p, \dots

$$E(x) = 0 \cdot p + 1 \cdot qp + 2q^2p + 3q^3p + \dots = pq(1 + 2q + 3q^2 + \dots)$$

$$= pq(1-q)^{-2} = pq \frac{1}{p^2} = \frac{q}{p} = \frac{1-p}{p} = \frac{1}{p} - 1.$$

Example 4. Show by an example that the mathematical expectation need not be finite.

Sol. Let x be the discontinuous random variable having values $0!, 1!, 2!, 3!, \dots, \infty$ and

let the probability law for it be $\frac{e^{-1}}{x!}$

$$E(x!) = 0! \cdot \frac{e^{-1}}{0!} + 1! \cdot \frac{e^{-1}}{1!} + 2! \cdot \frac{e^{-1}}{2!} + 3! \cdot \frac{e^{-1}}{3!} + \dots$$

$$= e^{-1} + e^{-1} + e^{-1} + e^{-1} + \dots$$

$$= e^{-1}(1 + 1 + \dots \infty)$$

$$= e^{-1} \sum_{0}^{\infty} 1, \text{ which is not finite.}$$

Example 5. Show that if x takes the values $x_n = (-1)^n \cdot 2^n \cdot n^{-1}$ for $n = 1, 2, \dots$ with probabilities $p_n = 2^{-n}$ then $E(x) = -\log 2$.

Sol. Here

$$\begin{aligned} E(x) &= \sum_{n=1}^{\infty} 2^{-n} \cdot (-1)^n \cdot 2^n \cdot n^{-1} \\ &= \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \dots \\ &= -\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots\right) = -\log 2. \end{aligned}$$

Example 6. What is the mathematical expectation of the sum of points on n dice?

Sol. Denoting by x_i the number of points on the i^{th} dice, the sum of points on n dice is

$$x_1 + x_2 + \dots + x_n$$

$$E(x_1 + x_2 + \dots + x_n) = E(x_1) + E(x_2) + \dots + E(x_n)$$

∴

But for every single dice $E(x_i) = \frac{7}{2}$, $i = 1, 2, 3, \dots, n$

$$\therefore E(x_1 + x_2 + \dots + x_n) = \frac{7}{2} + \frac{7}{2} + \dots \text{ } n \text{ terms} = \frac{7}{2} n.$$

Example 7. Find the expected value of the product of points on n dice.

Sol. Denoting by x_i the number of points on the i^{th} dice the product of points on n dice is

$$x_1, x_2, x_3, \dots, x_n$$

For every single dice $E(x_i) = \frac{7}{2}$, $i = 1, 2, 3, \dots, n$

$$\therefore E(x_1 x_2 \dots x_n) = E(x_1) \cdot E(x_2) \dots E(x_n) = \frac{7}{2} \cdot \frac{7}{2} \cdot \frac{7}{2} \dots n \text{ factors} = \left(\frac{7}{2}\right)^n.$$

Example 8. There is a series of n independent trials, where p_i is the probability for i^{th}

trial. Show that the mean number of successes is $\sum_{i=1}^n p_i$.

Sol. Let a variable x_i be associated with i^{th} trial, which has the value 1 in case of success and the value 0 in case of a failure. If m be the number of successes in n trials then

$$m = x_1 + x_2 + \dots + x_n$$

$$\therefore E(m) = E(x_1 + x_2 + \dots + x_n) = E(x_1) + E(x_2) + \dots + E(x_n)$$

Since x_i can take the values 1 or 0 with probabilities p_i or $1 - p_i$, we have

$$E(x_i) = 1, p_i + 0(1 - p_i) = p_i$$

$$\therefore E(m) = E(x_1) + E(x_2) + \dots + E(x_n) = p_1 + p_2 + \dots + p_n.$$

Example 9. Find the variance of the number of successes in a series of n independent trials in which the probability of success in the i^{th} trial is p_i .

Sol. Here

$$m = x_1 + x_2 + \dots + x_n$$

$$\therefore \text{var}(m) = \text{var}(x_1 + x_2 + \dots + x_n) = \text{var}(x_1) + \text{var}(x_2) + \dots + \text{var}(x_n)$$

If there is success, $x_i^2 = 1$, if failure, $x_i^2 = 0$

$$\therefore E(x_i^2) = 1 \cdot p_i + 0 \cdot (1 - p_i) = p_i.$$

Here the variance of the number of successes in the i^{th} trial is given by

$$\begin{aligned}\text{var}(x_i) &= E[x_i - E(x_i)]^2 \\ &= E(x_i)^2 - 2x_i E(x_i) + \{E(x_i)\}^2 = E(x_i)^2 - \{E(x_i)\}^2 \\ &= p_i - p_i^2 = p_i(1 - p_i) = p_i q_i\end{aligned}$$

$$\therefore \text{var}(m) = \sum_{i=1}^n p_i q_i = p_1 q_1 + p_2 q_2 + \dots + p_n q_n.$$

Corollary 1. Show that $\text{var}(m) \leq \frac{1}{4}$.

$$\text{Since } pq = p(1-p) = p - p^2 - \frac{1}{4} + \frac{1}{4} = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4}, \text{ or } \left(p - \frac{1}{2}\right)^2 \text{ is positive.}$$

Corollary 2. If $p_1 = p_2 = \dots = p_n = p$, then $\text{var}(m) = pq + pq + \dots + pq = npq$.

Example 10. From a point on the circumference of a circle of radius a , a chord is drawn in a random direction. Show that the expected value of the length of the chord is $4a/\pi$ and that the variance of the length is $2a^2(1 - 8/\pi^2)$. Also show that the chance is $\frac{1}{3}$ and the length of the chord will exceed the length of the side of an equilateral triangle inscribed in the circle.

Sol. Let C be the centre and A be any point. Let AB be any random chord.

Let $\angle DAB = \theta$. Then θ lies from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

$$dF(\theta) = f(\theta) d\theta = \frac{d\theta}{\frac{1}{2}\pi - \left(-\frac{1}{2}\pi\right)} = \frac{d\theta}{\pi}, \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

If a be the radius, $AB = 2a \cos \theta$

$$\therefore E(AB) = \int_{-\pi/2}^{\pi/2} AB \cdot f(\theta) d\theta = \frac{2a}{\pi} \int_{-\pi/2}^{\pi/2} \cos \theta d\theta = \frac{4a}{\pi}.$$

$$E[(AB)^2] = \int_{-\pi/2}^{\pi/2} (AB)^2 f(\theta) d\theta = \frac{4a^2}{\pi} \int_{-\pi/2}^{\pi/2} \cos^2 \theta d\theta = 2a^2.$$

$$\text{var}(AB) = E[(AB)^2] - \{E(AB)\}^2 = 2a^2 - 16a^2/\pi^2 = 2a^2(1 - 8/\pi^2).$$

Now as the inscribed equilateral triangle is $a\sqrt{3}$, therefore

$$P(AB > a\sqrt{3}) = P(2a \cos \theta > a\sqrt{3}) = P(\cos \theta > \sqrt{3}/2)$$

$$= P\left(|\theta| < \frac{1}{6}\pi\right) = P\left(-\frac{1}{6}\pi < \theta < \frac{1}{6}\pi\right)$$

$$= \int_{-\pi/6}^{\pi/2} f(|\theta|) d\theta = \frac{1}{\pi} \int_{-\pi/6}^{\pi/2} 1 \cdot d\theta = \frac{1}{\pi} \cdot \frac{\pi}{3} = \frac{1}{3}.$$

TEST YOUR KNOWLEDGE

If x is a variate and a is a constant, show that:

1. $E(a) = a$.
2. $E(ax) = aE(x)$.
3. $\text{var}(ax + b) = a^2 \cdot \text{var}(x)$.
4. $\text{cov}(ax, by) = ab \text{ cov}(x, y)$.
5. $\text{cov}(x + a, y - b) = \text{cov}(x, y)$.
6. $\text{cov}(x, x) = \text{var}(x)$.

7. Prove that:

$$\text{cov}(ax + by, cx + dy) = ac \text{ var}(x) + bd \text{ var}(y) + (ad + bc) \text{ cov}(x, y).$$

8. If x, y, z are random variables each with expectation 10 and variance 1, 4 and 9 respectively;

$$r(x, y) = 0, r(y, z) = \frac{1}{4}, r(x, z) = \frac{1}{4} \text{ obtain}$$

- (a) $E(x + y - 2z)$
- (b) $\text{cov}(x + z, y + z)$
- (c) $V(x - 2z)$
- (d) $\text{cov}(3x, 5z)$.

9. Two unbiased dice are thrown. Find the expected value of the sum of numbers of points drawn.
10. A coin is tossed until a head appears. What is the expectation of the number of tosses?
11. In a lottery m tickets are drawn out of n tickets numbered from 1 to n . What is the expectation and the variance of the sum of the numbers drawn?
12. Three urns contain respectively 3 green and 2 white balls, 5 green and 6 white balls and 2 green and 4 white balls. One ball is drawn from each urn. Find the expected number of white balls drawn out.

Hint. $0\left(\frac{3}{5} \cdot \frac{5}{11} \cdot \frac{2}{6}\right) + 1 \cdot \left(\frac{2}{5} + \frac{2}{5} \cdot \frac{2}{6} + \frac{3}{5} \cdot \frac{5}{11} \cdot \frac{4}{6} + \frac{3}{5} \cdot \frac{6}{11} \cdot \frac{2}{6}\right)$

$$+ 2\left(\frac{2}{5} \cdot \frac{6}{11} \cdot \frac{2}{6} + \frac{2}{5} \cdot \frac{5}{11} \cdot \frac{4}{6} + \frac{3}{5} \cdot \frac{6}{11} \cdot \frac{4}{6}\right) + 3 \cdot \left(\frac{2}{5} \cdot \frac{6}{11} \cdot \frac{4}{6}\right) = \frac{266}{165}$$

13. Show that the expectation of the number of failures preceding the first success in an indefinite series of independent trials with constant probability p of success, is

$$qp + 2q^2p + 3q^3p + \dots = \frac{pq}{(1-q)^2} = \frac{q}{p} = \frac{1-p}{p}.$$

14. Define mathematical expectation. Prove that mathematical expectation of the sum of two discrete random variables is equal to the sum of their mathematical expectations.

15. x denotes the profit that a man can make in business. He may earn ₹2,800 with probability $\frac{1}{2}$, he may lose ₹5,000 with probability $\frac{3}{10}$ and he may neither lose nor gain with probability $\frac{1}{5}$. Show that the mathematical expectation is -100.
16. If x is a random variable for which $E(x) = 10$ and $\text{var}(x) = 25$. Find the positive values of a and b such that $y = ax - b$ has expectation 0 and variance 1.

Answers

9. 7

10. 2

12. $\frac{266}{165}$

11. $\frac{1}{2}m(n+1); \frac{1}{12}m(n+1)(n-2)$

16. $a = \frac{1}{5}, b = 2$.

3.49 THEORETICAL PROBABILITY DISTRIBUTIONS

Generally, frequency distribution are formed from the observed or experimental data. However, frequency distribution of certain populations can be deduced mathematically by fitting theoretical probability distribution under certain assumptions.

Frequency distributions can be classified under two heads:

- (i) Observed Frequency Distributions.
- (ii) Theoretical or Expected Frequency Distributions.

Observed frequency distributions are based on actual observation and experimentation.

If certain hypothesis is assumed, it is sometimes possible to derive mathematically what

the frequency distribution of certain universe should be. Such distributions are called **Theoretical Distributions**.

Theoretical probability distributions are of two types:

(i) **Discrete probability distribution.** Binomial, poisson, geometric, negative binomial, hypergeometric, multinomial, multivariate hypergeometric distributions.

(ii) **Continuous probability distributions.**

Uniform, normal Gamma, exponential, χ^2 , Beta, bivariate normal, t , F-distributions.

Here, we will study three important theoretical probability distributions:

1. Binomial Distribution (or Bernoulli's Distribution)
2. Poisson's Distribution
3. Normal Distribution.

3.50 BINOMIAL PROBABILITY DISTRIBUTION

[G.B.T.U. 2010, 2013]

It was discovered by a Swiss Mathematician Jacob James Bernoulli in the year 1700.

This distribution is concerned with trials of a repetitive nature in which only the occurrence or non-occurrence, success or failure, acceptance or rejection, yes or no of a particular event is of interest.

For convenience, we shall call the occurrence of the event 'a success' and its non-occurrence 'a failure'.

Let there be n independent trials in an experiment. Let a random variable X denote the number of successes in these n trials. Let p be the probability of a success and q that of a failure in a single trial so that $p + q = 1$. Let the trials be independent and p be constant for every trial.

Let us find the probability of r successes in n trials.

r successes can be obtained in n trials in ${}^n C_r$ ways.

$$\begin{aligned} \therefore P(X = r) &= {}^n C_r \underbrace{P(S S S \dots S)}_{r \text{ times}} \quad \underbrace{F F F \dots F}_{(n-r) \text{ times}} \\ &= {}^n C_r \underbrace{P(S) P(S) \dots P(S)}_{r \text{ factors}} \quad \underbrace{P(F) P(F) \dots P(F)}_{(n-r) \text{ factors}} \\ &= {}^n C_r \underbrace{p p p \dots p}_{r \text{ factors}} \quad \underbrace{q q q \dots q}_{(n-r) \text{ factors}} \\ &= {}^n C_r p^r q^{n-r} \end{aligned} \quad \dots(1)$$

Hence $P(X = r) = {}^n C_r p^r q^{n-r}$ where $p + q = 1$ and $r = 0, 1, 2, \dots, n$.

The distribution (1) is called the *binomial probability distribution* and X is called the *binomial variate*.

Note 1. $P(X = r)$ is usually written as $P(r)$.

Note 2. The successive probabilities $P(r)$ in (1) for $r = 0, 1, 2, \dots, n$ are

$${}^n C_0 q^n, {}^n C_1 q^{n-1} p, {}^n C_2 q^{n-2} p^2, \dots, {}^n C_n p^n$$

which are the successive terms of the binomial expansion of $(q + p)^n$. That is why this distribution is called "binomial" distribution.

Note 3. n and p occurring in the binomial distribution are called the *parameters* of the distribution.

Note 4. In a binomial distribution:

- (i) n , the number of trials is finite.
- (ii) each trial has only two possible outcomes usually called success and failure.
- (iii) all the trials are independent.
- (iv) p (and hence q) is constant for all the trials.

3.51 RECURRENCE OR RECURSION FORMULA FOR THE BINOMIAL DISTRIBUTION

In a binomial distribution,

$$\begin{aligned} P(r) &= {}^n C_r q^{n-r} p^r = \frac{n!}{(n-r)! r!} q^{n-r} p^r \\ P(r+1) &= {}^n C_{r+1} q^{n-r-1} p^{r+1} = \frac{n!}{(n-r-1)! (r+1)!} q^{n-r-1} p^{r+1} \\ \therefore \frac{P(r+1)}{P(r)} &= \frac{(n-r)!}{(n-r-1)!} \times \frac{r!}{(r+1)!} \times \frac{p}{q} \\ &= \frac{(n-r) \times (n-r-1)!}{(n-r-1)!} \times \frac{r!}{(r+1) \times r!} \times \frac{p}{q} = \left(\frac{n-r}{r+1} \right) \cdot \frac{p}{q} \\ \Rightarrow P(r+1) &= \frac{n-r}{r+1} \cdot \frac{p}{q} P(r) \end{aligned}$$

which is the required recurrence formula. Applying this formula successively, we can find $P(1)$, $P(2)$, $P(3)$, ..., if $P(0)$ is known.

3.52 MEAN AND VARIANCE OF THE BINOMIAL DISTRIBUTION

[U.P.T.U. 2008, G.B.T.U. 2012]

For the binomial distribution, $P(r) = {}^n C_r q^{n-r} p^r$

$$\begin{aligned} \text{Mean } \mu &= \sum_{r=0}^n r P(r) = \sum_{r=0}^n r \cdot {}^n C_r q^{n-r} p^r \\ &= 0 + 1 \cdot {}^n C_1 q^{n-1} p + 2 \cdot {}^n C_2 q^{n-2} p^2 + 3 \cdot {}^n C_3 q^{n-3} p^3 + \dots + n \cdot {}^n C_n p^n \\ &= nq^{n-1} p + 2 \cdot \frac{n(n-1)}{2 \cdot 1} q^{n-2} p^2 + 3 \cdot \frac{n(n-1)(n-2)}{3 \cdot 2 \cdot 1} q^{n-3} p^3 + \dots + n \cdot p^n \\ &= nq^{n-1} p + n(n-1) q^{n-2} p^2 + \frac{n(n-1)(n-2)}{2 \cdot 1} q^{n-3} p^3 + \dots + np^n \\ &= np[{}^{n-1} C_0 q^{n-1} + {}^{n-1} C_1 q^{n-2} p + {}^{n-1} C_2 q^{n-3} p^2 + \dots + {}^{n-1} C_{n-1} p^{n-1}] \\ &= np(q+p)^{n-1} = np \quad (\because p+q=1) \end{aligned}$$

Hence the mean of the binomial distribution is np .

$$\begin{aligned} \text{Variance } \sigma^2 &= \sum_{r=0}^n r^2 P(r) - \mu^2 = \sum_{r=0}^n [r + r(r-1)] P(r) - \mu^2 \\ &= \sum_{r=0}^n r P(r) + \sum_{r=0}^n r(r-1) P(r) - \mu^2 = \mu + \sum_{r=2}^n r(r-1) {}^n C_r q^{n-r} p^r - \mu^2 \end{aligned}$$

(since the contribution due to $r=0$ and $r=1$ is zero)

$$\begin{aligned}
&= \mu + [2 \cdot 1 \cdot {}^nC_2 q^{n-2} p^2 + 3 \cdot 2 \cdot {}^nC_3 q^{n-3} p^3 + \dots + n(n-1) {}^nC_n p^n] - \mu^2 \\
&= \mu + \left[2 \cdot 1 \cdot \frac{n(n-1)}{2 \cdot 1} q^{n-2} p^2 + 3 \cdot 2 \cdot \frac{n(n-1)(n-2)}{3 \cdot 2 \cdot 1} q^{n-3} p^3 + \dots + n(n-1) p^n \right] - \mu^2 \\
&= \mu + [n(n-1)q^{n-2}p^2 + n(n-1)(n-2)q^{n-3}p^3 + \dots + n(n-1)p^n] - \mu^2 \\
&= \mu + n(n-1)p^2[q^{n-2} + (n-2)q^{n-3}p + \dots + p^{n-2}] - \mu^2 \\
&= \mu + n(n-1)p^2[{}^{n-2}C_0 q^{n-2} + {}^{n-2}C_1 q^{n-3}p + \dots + {}^{n-2}C_{n-2} p^{n-2}] - \mu^2 \\
&= \mu + n(n-1)p^2(q+p)^{n-2} - \mu^2 = \mu + n(n-1)p^2 - \mu^2 & [\because q+p=1] \\
&= np + n(n-1)p^2 - n^2 p^2 = np[1-p] = npq. & [\because \mu=np]
\end{aligned}$$

Hence the variance of the binomial distribution is npq .

Standard deviation of the binomial distribution is \sqrt{npq} .

3.53 MOMENT GENERATING FUNCTION OF BINOMIAL DISTRIBUTION

1. About origin

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^n e^{tx} {}^nC_x p^x q^{n-x} = \sum_{x=0}^n {}^nC_x (pe^t)^x q^{n-x} = (q + pe^t)^n$$

2. About mean

[G.B.T.U. 2012, U.P.T.U. 2008, 2015]

$$\begin{aligned}
M_{x-np}(t) &= E[e^{t(x-np)}] \\
&= e^{-npt} E(e^{tx}) = e^{-npt} M_x(t) = e^{-npt} (q + pe^t)^n \\
&= (qe^{-pt} + pe^{t-pt})^n = (qe^{-pt} + pe^{qt})^n & |\because 1-p=q
\end{aligned}$$

3.54 MOMENTS ABOUT MEAN OF BINOMIAL DISTRIBUTION

$$\begin{aligned}
M_{x-np}(t) &= (qe^{-pt} + pe^{qt})^n \\
&= \left[q \left(1 - pt + \frac{p^2 t^2}{2!} - \frac{p^3 t^3}{3!} + \dots \right) + p \left(1 + qt + \frac{q^2 t^2}{2!} + \frac{q^3 t^3}{3!} + \dots \right) \right]^n \\
&= \left[(q + p) + \frac{t^2}{2!} pq (q + p) + \frac{t^3}{3!} pq (q^2 - p^2) + \frac{t^4}{4!} pq (q^3 + p^3) + \dots \right]^n \\
&= \left[1 + \left\{ \frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq (q - p) + \frac{t^4}{4!} qp (1 - 3pq) + \dots \right\} \right]^n \\
&= \left[1 + {}^nC_1 \left\{ \frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq (q - p) + \frac{t^4}{4!} pq (1 - 3pq) + \dots \right\} \right. \\
&\quad \left. + {}^nC_2 \left\{ \frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq (q - p) + \dots \right\}^2 + \dots \right]
\end{aligned}$$

Now,

$$\mu_2 = \text{coefficient of } \frac{t^2}{2!} = npq$$

$$\mu_3 = \text{coefficient of } \frac{t^3}{3!} = npq(q-p)$$

$$\begin{aligned}\mu_4 &= \text{coefficient of } \frac{t^4}{4!} = npq(1-3pq) + 3n(n-1)p^2q^2 \\ &= 3n^2p^2q^2 + npq(1-6pq)\end{aligned}$$

Hence,

$$\beta_1 = \frac{\mu_3^2}{\mu_2^2} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\therefore \gamma_1 = \frac{1-2p}{\sqrt{npq}}$$

$$\text{and } \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

$$\therefore \gamma_2 = \frac{1-6pq}{npq}$$

Note 1. $\gamma_1 = \frac{1-2p}{\sqrt{npq}}$ gives a **measure of skewness** of the binomial distribution. If $p < \frac{1}{2}$, skewness is positive, if $p > \frac{1}{2}$, skewness is negative and if $p = \frac{1}{2}$, it is zero.

$\beta_2 = 3 + \frac{1-6pq}{npq}$ gives a **measure of the kurtosis** of the binomial distribution.

Note 2. If n independent trials constitute one experiment and this experiment is repeated N times then the frequency of r successes is $N \cdot {}^n C_r p^r q^{n-r}$.

3.55 APPLICATIONS OF BINOMIAL DISTRIBUTION

1. In problems concerning no. of defectives in a sample production line.
2. In estimation of reliability of systems.
3. No. of rounds fired from a gun hitting a target.
4. In Radar detection.

EXAMPLES

Example 1. (i) Comment on the following statement:

For a Binomial distribution, mean is 6 and variance is 9.

(ii) A die is tossed thrice. A success is getting 1 or 6 on a toss. Find the mean and variance of the number of success.

$$\text{Sol. (i)} \quad \mu = np = 6 \quad \dots(1)$$

$$\sigma^2 = npq = 9 \quad \dots(2)$$

Dividing (2) by (1), we get

$$q = \frac{9}{6} = 1.5$$

which is impossible as $0 \leq q \leq 1$

\therefore The above statement is **False**.

(ii) Prob. of getting success (1 or 6) on a toss = $\frac{2}{6} = \frac{1}{3} = p$

$$\therefore q = 1 - \frac{1}{3} = \frac{2}{3}$$

No. of tosses of a die, $n = 3$

$$(i) \text{ Mean} = np = 3\left(\frac{1}{3}\right) = 1. \quad (ii) \text{ Variance} = npq = (3)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{3}.$$

Example 2. If 10% of the bolts produced by a machine are defective, determine the probability that out of 10 bolts chosen at random

(i) 1 (ii) None (iii) at most 2 bolts will be defective.

Sol. Here, $p(\text{defective}) = \frac{10}{100} = \frac{1}{10}$ (given)

$$\therefore q(\text{non-defective}) = 1 - \frac{1}{10} = \frac{9}{10}$$

Also, $n = 10$, (n is no. of bolts chosen). (given)

The probability of r defective bolts out of n bolts chosen at random is given by

$$P(r) = {}^nC_r p^r q^{n-r} \quad \dots(1)$$

(i) Here $r = 1$,

$$\therefore P(1) = {}^{10}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{10-1} \quad | \text{ Using (1)}$$

$$= 10 \left(\frac{1}{10}\right) \left(\frac{9}{10}\right)^9 = (.9)^9 = 0.3874 \quad \dots(2)$$

(ii) Here $r = 0$

$$\therefore P(0) = {}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10-0} = \left(\frac{9}{10}\right)^{10} = 0.3486 \quad \dots(3) | \text{ Using (1)}$$

$$(iii) \text{ Prob. that at most 2 bolts will be defective} = P(r \leq 2) = P(0) + P(1) + P(2) \quad \dots(4)$$

Now, $P(2) = {}^{10}C_2 \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{10-2} \quad | \text{ Using (1)}$

$$= 45 \left(\frac{1}{100}\right) (0.43046) = 0.1937$$

$$\therefore \text{From (4), Required Probability} = P(0) + P(1) + P(2) \\ = 0.3486 + 0.3874 + 0.1937 = 0.9297.$$

Example 3. A binomial variable X satisfies the relation $9P(X = 4) = P(X = 2)$ when $n = 6$. Find the value of the parameter p and $P(X = 1)$.

Sol. We know that

$$P(X = r) = {}^nC_r p^r q^{n-r} \quad \dots(1)$$

$$\therefore P(X = 4) = {}^6C_4 p^4 q^2 = 15p^4 q^2$$

and $P(X = 2) = {}^6C_2 p^2 q^4 = 15 p^2 q^4 \quad | \text{ Since } n = 6$

The given relation is

$$9 P(X = 4) = P(X = 2) \Rightarrow 9(15p^4 q^2) = 15p^2 q^4$$

$$\begin{aligned}
 \Rightarrow & 9p^2 = q^2 = (1-p)^2 & | \because p+q=1 \\
 \Rightarrow & 9p^2 = 1 + p^2 - 2p \\
 \Rightarrow & 8p^2 + 2p - 1 = 0 \quad \Rightarrow \quad (4p-1)(2p+1) = 0 \\
 \therefore & p = \frac{1}{4} & | \because p \text{ cannot be negative}
 \end{aligned}$$

Now, $P(X=1) = {}^6C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^5 = .3559.$ $| \because q = 1 - p = \frac{3}{4}$

Example 4. Fit a binomial distribution to the following frequency data:

$x :$	0	1	2	3	4
$f :$	30	62	46	10	2

Sol. The table is as follows:

x	f	fx
0	30	0
1	62	62
2	46	92
3	10	30
4	2	8
	$\Sigma f = 150$	$\Sigma fx = 192$

$$\begin{aligned}
 \text{Mean of observations} &= \frac{\Sigma fx}{\Sigma f} = \frac{192}{150} = 1.28 \\
 \Rightarrow & np = 1.28 \\
 \Rightarrow & 4p = 1.28 \quad (n \text{ is no. of trials}) \\
 \Rightarrow & p = 0.32 \\
 \therefore & q = 1 - p = 1 - 0.32 = 0.68 \\
 \text{Also,} & N = 150 \quad | \because N = \Sigma f
 \end{aligned}$$

Hence the binomial distribution is $= N(q+p)^n = 150 (0.68 + 0.32)^4.$

Example 5. A student is given a true-false examination with 8 questions. If he corrects at least 7 questions, he passes the examination. Find the probability that he will pass given that he guesses all questions.

Sol. Here, $n = \text{no. of questions asked} = 8$

$$p = \frac{1}{2}, q = \frac{1}{2} \quad | \text{ Since the question can either be true or false}$$

Probability that he will pass

$$\begin{aligned}
 &= P(r \geq 7) = P(7) + P(8) \\
 &= {}^8C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{8-7} + {}^8C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{8-8} = 8 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^1 + 1 \cdot \left(\frac{1}{2}\right)^8 \\
 &= \left(\frac{1}{2}\right)^8 (8+1) = \frac{9}{256} = .03516.
 \end{aligned}$$

Example 6. During war, 1 ship out of 9 was sunk on an average in making a certain voyage. What was the probability that exactly 3 out of a convoy of 6 ships would arrive safely?

Sol. p , the probability of a ship arriving safely $= 1 - \frac{1}{9} = \frac{8}{9}$; $q = \frac{1}{9}$, $n = 6$

The probability that exactly 3 ships arrive safely $= P(r = 3) = {}^6C_3 \left(\frac{1}{9}\right)^3 \left(\frac{8}{9}\right)^3 = \frac{10240}{9^6}$.

Example 7. A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that dacoit is still alive?

Sol. Here n = no. of bullets fired $= 6$, $p = 0.6$, $q = 1 - p = 0.4$

Probability that dacoit is still alive (not killed)

$$= P(r = 0) = {}^nC_0 p^0 q^{n-0} = {}^6C_0 (.6)^0 (.4)^6 = (.4)^6 = .004096.$$

Example 8. If the probability of hitting a target is 10% and 10 shots are fired independently. What is the probability that the target will be hit at least once?

Sol. Here, $p = \frac{10}{100} = \frac{1}{10}$, $q = 1 - p = 1 - \frac{1}{10} = \frac{9}{10}$, $n = 10$

Probability that the target will be hit at least once

$$\begin{aligned} &= P(r \geq 1) = 1 - P(r = 0) \\ &= 1 - [{}^nC_0 p^0 q^n] = 1 - \left[{}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10} \right] = 0.6513. \end{aligned}$$

Example 9. Out of 800 families with 4 children each, how many families would be expected to have (i) 2 boys and 2 girls (ii) at least one boy (iii) no girl (iv) atmost two girls? Assume equal probabilities for boys and girls. (U.P.T.U. 2014)

Sol. Since probabilities for boys and girls are equal,

$$p = \text{probability of having a boy} = \frac{1}{2}; q = \text{probability of having a girl} = \frac{1}{2}$$

$$n = 4 \quad N = 800$$

(i) The expected number of families having 2 boys and 2 girls

$$= 800 \cdot {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 800 \times 6 \times \frac{1}{16} = 300.$$

(ii) The expected number of families having at least one boy

$$\begin{aligned} &= 800 \left[{}^4C_1 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right) + {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 + {}^4C_3 \left(\frac{1}{2}\right) \left(\frac{1}{2}\right)^3 + {}^4C_4 \left(\frac{1}{2}\right)^4 \right] \\ &= 800 \times \frac{1}{16} [4 + 6 + 4 + 1] = 750 \end{aligned}$$

(iii) The expected number of families having no girl i.e., having 4 boys

$$= 800 \cdot {}^4C_4 \left(\frac{1}{2}\right)^4 = 50.$$

(iv) The expected number of families having atmost two girls i.e., having at least 2 boy

$$= 800 \left[{}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 + {}^4C_3 \left(\frac{1}{2}\right) \left(\frac{1}{2}\right)^3 + {}^4C_4 \left(\frac{1}{2}\right)^4 \right] = 800 \times \frac{1}{16} [6 + 4 + 1] = 550.$$

Example 10. Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six?

$$\text{Sol. } p = \text{the chance of getting 5 or 6 with one die} = \frac{2}{6} = \frac{1}{3}$$

$$q = 1 - \frac{1}{3} = \frac{2}{3}, n = 6, N = 729$$

since dice are in sets of 6 and there are 729 sets.

The expected number of times at least three dice showing five or six

$$\begin{aligned} &= N \cdot P(r \geq 3) \\ &= 729 [P(3) + P(4) + P(5) + P(6)] \\ &= 729 \left[{}^6C_3 \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^3 + {}^6C_4 \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^4 + {}^6C_5 \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^5 + {}^6C_6 \left(\frac{1}{3}\right)^6 \right] \\ &= \frac{729}{3^6} [160 + 60 + 12 + 1] = 233. \end{aligned}$$

Example 11. The probability of a man hitting a target is $\frac{1}{3}$. How many times must he fire so that the probability of his hitting the target at least once is more than 90%?

$$\text{Sol. } p = \frac{1}{3}$$

The probability of not hitting the target in n trials is q^n .

Therefore, to find the smallest n for which the probability of hitting at least once is more than 90%, we have

$$\begin{aligned} 1 - q^n &> 0.9 \\ \Rightarrow 1 - \left(\frac{2}{3}\right)^n &> 0.9 \\ \Rightarrow \left(\frac{2}{3}\right)^n &< 0.1 \end{aligned}$$

The smallest n for which the above inequality holds true is 6 hence he must fire 6 times.

Example 12. In a bombing action, there is 50% chance that any bomb will strike the target. Two direct hits are needed to destroy the target completely. How many bombs are required to be dropped to give a 99% chance or better of completely destroying the target?

$$\text{Sol. We have, } p = \frac{50}{100} = \frac{1}{2}$$

Since the probability must be greater than 0.99, if n bombs are dropped, we have

$$\begin{aligned} {}^nC_2 \left(\frac{1}{2}\right)^n + {}^nC_3 \left(\frac{1}{2}\right)^n + {}^nC_4 \left(\frac{1}{2}\right)^n + \dots + {}^nC_n \left(\frac{1}{2}\right)^n &\geq 0.99 \\ \Rightarrow \left(\frac{1}{2}\right)^n [{}^nC_2 + {}^nC_3 + {}^nC_4 + \dots + {}^nC_n] &\geq 0.99 \\ \frac{2^n - n - 1}{2^n} &\geq 0.99 \end{aligned}$$

$$\begin{aligned}
 &\Rightarrow 1 - \frac{1+n}{2^n} \geq 0.99 \\
 &\Rightarrow \frac{1+n}{2^n} \leq 0.01 \\
 &\Rightarrow 2^n \geq 100n + 100
 \end{aligned}$$

By trial, $n = 11$ satisfies the inequality.

Hence 11 bombs are required to be dropped.

ASSIGNMENT

1. (i) Ten coins are tossed simultaneously. Find the probability of getting at least seven heads.
 (ii) A die is thrown five times. If getting an odd number is a success, find the probability of getting at least four successes. (M.T.U. 2012)
2. (a) The probability of any ship of a company being destroyed on a certain voyage is 0.02. The company owns 6 ships for the voyage. What is the probability of:
 (i) losing one ship (ii) losing atmost two ships (iii) losing none?
 (b) Assume that on the average one telephone number out of fifteen called between 2 P.M. and 3 P.M. on week-days is busy. What is the probability that if 6 randomly selected telephone numbers are called (i) not more than 3 (ii) at least 3 of them will be busy?
3. (i) The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of six workers chosen at random, four or more will suffer from the disease?
 (ii) The probability that a man aged 60 will live to be 70 is 0.65. What is the probability that out of 10 men, now 60, at least 7 will live to be 70?
4. (i) If the mean of a binomial distribution is 3 and the variance is $\frac{3}{2}$, find the probability of obtaining at least 4 successes.
 (ii) In a binomial distribution, for $n = 5$ if $P(x = 1) = 0.4096$ and $P(x = 2) = 0.2048$, then find the value of p .
 (iii) The sum and product of the mean and variance of a binomial distribution are $\frac{25}{3}$ and $\frac{50}{3}$ respectively. Find the distribution. (U.P.T.U. 2007)
 (iv) If the probability of a defective bolt is 0.1, find (a) The mean (b) The standard deviation for the distribution in a total of 400 bolts.
 (v) If the moment generating function of a random variable X is $\left(\frac{1}{3} + \frac{2}{3} e^t\right)^5$, find $P(X = 2)$.
5. (a) The probability that a bomb dropped from a plane will strike the target is $\frac{1}{5}$. If six bombs are dropped, find the probability that (i) exactly two will strike the target, (ii) at least two will strike the target.
 (b) Four persons in a group of 20 are graduates. If 4 persons are selected at random from 20, find the probability that
 (i) all are graduates (ii) at least one is a graduate.
6. A bag contains 5 white, 7 red and 8 black balls. If four balls are drawn, one by one, with replacement, what is the probability that
 (i) none is white (ii) all are white
 (iii) at least one is white (iv) only 2 are white?

7. (i) In a hurdle race, a player has to cross 10 hurdles. The probability that he will clear each hurdle is $\frac{5}{6}$. What is the probability that he will knock down fewer than 2 hurdles ?
(ii) If on an average one ship in every ten is wrecked, find the probability that out of 5 ships expected to arrive, 4 at least will arrive safely.

8. The prob. that a bulb produced by a factory will fuse after a use of 150 days is 0.05. Find the prob. that out of 5 such bulbs [M.T.U. (B. Pharma) 2011]
(i) None
(ii) Atmost one
(iii) More than one
(iv) At least one will fuse after 150 days of use.

9. Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. What is the probability that
(i) all the five cards are spades (ii) only three are spades (iii) none is spade?

10. Manish takes a step forward with probability 0.4 and backward with probability 0.6. Find the probability that at the end of 11 steps, he is one step away from the starting point.
[Hint. Manish will take either 6 steps forward and 5 backward or 5 steps forward and 6 backward].

11. (a) In 800 families with 5 children each, how many families would be expected to have (i) 3 boys and 2 girls, (ii) 2 boys and 3 girls, (iii) no girl (iv) at the most two girls. (Assume probabilities for boys and girls to be equal.)
(b) Out of 800 families with 5 children each, how many would you expect to have (i) 3 boys (ii) 5 girls (iii) either 2 or 3 boys? Assume equal probabilities for boys and girls.
(M.T.U. 2012)

(c) Out of 320 families with 5 children each, what percentage would be expected to have (i) 2 boys and 3 girls (ii) at least one boy? Assume equal probability for boys and girls.
(G.B.T.U. 2011)

12. In 100 sets of ten tosses of an unbiased coin, in how many cases do you expect to get
(i) 7 heads and 3 tails (ii) at least 7 heads?

13. The following data are the number of seeds germinating out of 10 on damp filter for 80 sets of seeds. Fit a binomial distribution to this data:

$x:$	0	1	2	3	4	5	6	7	8	9	10	Total
$f:$	6	20	28	12	8	6	0	0	0	0	0	80

[Hint. Here $n = 10$, $N = 80$, Mean = $\frac{\sum fx}{\sum f} = \frac{2.175}{80} = 2.175$ $\therefore np = 2.175$ etc.]

14. Fit a binomial distribution for the following data and compare the theoretical frequencies with the actual ones :

$x:$	0	1	2	3	4	5
$f:$	2	14	20	34	22	8

15. Fit a binomial distribution to the data given in the following table:
(i) $x:$ 0 1 2 3 4
 $f:$ 24 41 28 5 2 (M.T.U. 2012)
(ii) $x:$ 0 1 3 4
 $f:$ 28 62 10 4 (U.K.T.U. 2011)

16. (i) Assuming half the population of a town consumes chocolates so that the chance of an individual being consumer is $\frac{1}{2}$ and that 100 investigators each take 10 individuals to see whether they are consumers, how many investigators would you expect to report that 3 people or less were consumers? (U.P.T.U. 2006)

- (ii) Assuming that 20% of the population of a city are literate, so that the chance of an individual being literate is $\frac{1}{5}$ and assuming that 100 investigators each take 10 individuals to see whether they are literate, how many investigators would you expect to report 3 or less were literate?

17. Following results were obtained when 100 batches of seeds were allowed to germinate on damp filter paper in a laboratory : $\beta_1 = \frac{1}{15}$, $\beta_2 = \frac{89}{30}$. Determine the Binomial distribution. Calculate the expected frequency for $x = 8$ assuming $p > q$.

18. A coffee connoisseur claims that he can distinguish between a cup of instant coffee and a cup of percolator coffee 75% of the time. It is agreed that his claim will be accepted if he correctly identifies at least 5 of the 6 cups. Find his chances of having the claim (i) accepted (ii) rejected when he does have the ability he claims.

19. A multiple-choice test consists of 8 questions with 3 answers to each question of which only one is correct. A student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4 and the third answer if he gets 5 or 6. To get a distinction, the student must secure at least 75% correct answers. If there is no negative marking, what is the probability that the student secures a distinction?

20. An irregular six-faced die is thrown and the expectation that in 10 throws, it will give five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws each, would you expect it to give no even number?

Answers

16. (i) 17

(ii) 88

$$17. 100 \left(\frac{1}{4} + \frac{3}{4} \right)^{20}, 0.075168752$$

18. (i) 0.534

(ii) 0.466

19. 0.0197

20. 1.

POISSON DISTRIBUTION

3.56 POISSON DISTRIBUTION AS A LIMITING CASE OF BINOMIAL DISTRIBUTION

[U.P.T.U. 2007; G.B.T.U. 2010, 2013; M.T.U. 2013, 2014]

Poisson distribution was discovered by S.D. Poisson in the year 1837.

If the parameters n and p of a binomial distribution are known, we can find the distribution. But in situations where n is very large and p is very small, application of binomial distribution is very labourious. However, if we assume that as $n \rightarrow \infty$ and $p \rightarrow 0$ such that np always remains finite, say λ , we get the Poisson approximation to the binomial distribution.

Now, for a binomial distribution

$$\begin{aligned}
 P(X = r) &= {}^n C_r q^{n-r} p^r \\
 &= \frac{n(n-1)(n-2) \dots (n-r+1)}{r!} \times (1-p)^{n-r} \times p^r \\
 &= \frac{n(n-1)(n-2) \dots (n-r+1)}{r!} \times \left(1 - \frac{\lambda}{n}\right)^{n-r} \times \left(\frac{\lambda}{n}\right)^r \quad | \text{ Since } np = \lambda \quad \therefore \quad p = \frac{\lambda}{n} \\
 &= \frac{\lambda^r}{r!} \times \frac{n(n-1)(n-2) \dots (n-r+1)}{n^r} \times \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^r} \\
 &= \frac{\lambda^r}{r!} \left(\frac{n}{n}\right) \left(\frac{n-1}{n}\right) \left(\frac{n-2}{n}\right) \dots \left(\frac{n-r+1}{n}\right) \times \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^r} \\
 &= \frac{\lambda^r}{r!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{r-1}{n}\right) \times \frac{\left[\left(1 - \frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}}{\left(1 - \frac{\lambda}{n}\right)^r}
 \end{aligned}$$

As $n \rightarrow \infty$, each of the $(r-1)$ factors

$$\left(1 - \frac{1}{n}\right), \left(1 - \frac{2}{n}\right), \dots, \left(1 - \frac{r-1}{n}\right) \quad \text{tends to 1. Also } \left(1 - \frac{\lambda}{n}\right)^r \text{ tends to 1.}$$

Since $\lim_{x \rightarrow \pm\infty} \left(1 + \frac{1}{x}\right)^x = e$, the Naperian base. $\therefore \left[\left(1 - \frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} \rightarrow e^{-\lambda}$ as $n \rightarrow \infty$

Hence in the limiting case when $n \rightarrow \infty$, we have

$$P(X = r) = \frac{e^{-\lambda} \cdot \lambda^r}{r!} \quad (r = 0, 1, 2, 3, \dots) \quad \dots(1)$$

where λ is a finite number $= np$.

(1) represents a probability distribution which is called the *Poisson probability distribution*.

Note 1. λ is called the parameter of the distribution.

Note 2. The sum of the probabilities $P(r)$ for $r = 0, 1, 2, 3, \dots$ is 1, since

$$\begin{aligned} P(0) + P(1) + P(2) + P(3) + \dots &= e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \frac{\lambda^3 e^{-\lambda}}{3!} + \dots \\ &= e^{-\lambda} \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots \right) = e^{-\lambda} \cdot e^\lambda = 1. \end{aligned}$$

3.57 RECURRENCE FORMULA FOR THE POISSON DISTRIBUTION

[U.P.T.U. (B. Pharma.) 2009]

$$\text{For Poisson distribution, } P(r) = \frac{e^{-\lambda} \lambda^r}{r!} \quad \text{and} \quad P(r+1) = \frac{e^{-\lambda} \lambda^{r+1}}{(r+1)!}$$

$$\therefore \frac{P(r+1)}{P(r)} = \frac{\lambda r!}{(r+1)!} = \frac{\lambda}{r+1}$$

$$\text{or} \quad P(r+1) = \frac{\lambda}{r+1} P(r), \quad r = 0, 1, 2, 3, \dots$$

This is called the *recurrence or recursion formula* for the Poisson distribution.

3.58 MEAN AND VARIANCE OF THE POISSON DISTRIBUTION

[U.P.T.U. 2006]

$$\text{For the Poisson distribution, } P(r) = \frac{e^{-\lambda} \lambda^r}{r!}$$

$$\begin{aligned} \text{Mean } \mu &= \sum_{r=0}^{\infty} r P(r) = \sum_{r=0}^{\infty} r \cdot \frac{e^{-\lambda} \lambda^r}{r!} \\ &= e^{-\lambda} \sum_{r=1}^{\infty} \frac{\lambda^r}{(r-1)!} = e^{-\lambda} \left(\lambda + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \dots \right) \\ &= \lambda e^{-\lambda} \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots \right) = \lambda e^{-\lambda} \cdot e^\lambda = \lambda \end{aligned}$$

Thus, the mean of the Poisson distribution is equal to the parameter λ .

$$\begin{aligned} \text{Variance } \sigma^2 &= \sum_{r=0}^{\infty} r^2 P(r) - \mu^2 = \sum_{r=0}^{\infty} r^2 \cdot \frac{\lambda^r e^{-\lambda}}{r!} - \lambda^2 = e^{-\lambda} \sum_{r=1}^{\infty} \frac{r^2 \lambda^r}{r!} - \lambda^2 \\ &= e^{-\lambda} \left[\frac{1^2 \cdot \lambda}{1!} + \frac{2^2 \cdot \lambda^2}{2!} + \frac{3^2 \cdot \lambda^3}{3!} + \frac{4^2 \cdot \lambda^4}{4!} + \dots \right] - \lambda^2 \end{aligned}$$

$$\begin{aligned}
&= \lambda e^{-\lambda} \left[1 + \frac{2\lambda^2}{1!} + \frac{3\lambda^2}{2!} + \frac{4\lambda^3}{3!} + \dots \right] - \lambda^2 \\
&= \lambda e^{-\lambda} \left[1 + \frac{(1+1)\lambda}{1!} + \frac{(1+2)\lambda^2}{2!} + \frac{(1+3)\lambda^3}{3!} + \dots \right] - \lambda^2 \\
&= \lambda e^{-\lambda} \left[\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots \right) + \left(\frac{\lambda}{1!} + \frac{2\lambda^2}{2!} + \frac{3\lambda^3}{3!} + \dots \right) \right] - \lambda^2 \\
&= \lambda e^{-\lambda} \left[e^\lambda + \lambda \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots \right) \right] - \lambda^2 \\
&= \lambda e^{-\lambda} [e^\lambda + \lambda e^\lambda] - \lambda^2 = \lambda e^{-\lambda} \cdot e^\lambda (1 + \lambda) - \lambda^2 = \lambda(1 + \lambda) - \lambda^2 = \lambda.
\end{aligned}$$

Hence, the variance of the Poisson distribution is also λ .

Thus, the mean and the variance of the Poisson distribution are each equal to the parameter λ .

Note 1. The mean and the variance of the Poisson distribution can also be derived from those of the binomial distribution in the limiting case when $n \rightarrow \infty$, $p \rightarrow 0$ and $np = \lambda$.

Mean of Binomial distribution is np .

$$\therefore \text{Mean of Poisson distribution} = \lim_{n \rightarrow \infty} np = \lim_{n \rightarrow \infty} \lambda = \lambda$$

Variance of Binomial distribution is $npq = np(1-p)$

$$\therefore \text{Variance of Poisson distribution} = \lim_{n \rightarrow \infty} np(1-p) = \lim_{n \rightarrow \infty} \lambda \left(1 - \frac{\lambda}{n} \right) = \lambda.$$

Note 2. For Poisson distribution, $\mu_3 = \lambda$ and $\mu_4 = 3\lambda^2 + \lambda$.

Coefficients of skewness and kurtosis are given by

$$\beta_1 = \frac{1}{\lambda} \text{ and } \gamma_1 = \frac{1}{\sqrt{\lambda}}. \text{ Also, } \beta_2 = 3 + \frac{1}{\lambda} \text{ and } \gamma_2 = \frac{1}{\lambda}$$

Hence Poisson distribution is always a skewed distribution.

Remark. While fitting the Poisson distribution to a given data, we round the figures to the nearest integer but it should be kept in mind that the total of the observed and the expected frequencies should be same.

3.59 MODE OF POISSON DISTRIBUTION

Let $P(x = r) = e^{-\lambda} \frac{\lambda^r}{r!}, r = 0, 1, 2, \dots, \infty$

The value of r which has a greater probability than any other value is the mode of the Poisson distribution. Thus r is mode if

$$P(X = r) \geq P(X = r + 1) \text{ and } P(X = r) \geq P(X = r - 1)$$

$$\Rightarrow \frac{e^{-\lambda} \cdot \lambda^r}{r!} \geq \frac{e^{-\lambda} \cdot \lambda^{r+1}}{(r+1)!} \quad \text{and} \quad \frac{e^{-\lambda} \cdot \lambda^r}{r!} \geq \frac{e^{-\lambda} \cdot \lambda^{r-1}}{(r-1)!}$$

$$\Rightarrow 1 \geq \frac{\lambda}{r+1} \quad \text{and} \quad \frac{\lambda}{r} \geq 1$$

$$\Rightarrow r \geq \lambda - 1 \quad \text{and} \quad r \leq \lambda \quad i.e., \quad \lambda - 1 \leq r \leq \lambda$$

Case I. If λ is a positive integer, there are two modes $\lambda - 1$ and λ .

Case II. If λ is not a positive integer, there is one mode and is the integral value between $\lambda - 1$ and λ .

3.60 APPLICATIONS OF POISSON DISTRIBUTION

This distribution is applied to problems concerning :

- (i) Arrival pattern of defective vehicles in a workshop.
- (ii) Patients in a hospitals.
- (iii) Telephone calls.
- (iv) Demand pattern for certain spare parts.
- (v) Number of fragments from a shell hitting a target.
- (vi) Emission of radioactive (α) particles.

EXAMPLES

Example 1. If the variance of the Poisson distribution is 2, find the probabilities for $r = 1, 2, 3, 4$ from the recurrence relation of the Poisson distribution. Also find $P(r \geq 4)$.

(M.T.U. 2013)

Sol. λ , the parameter of Poisson distribution = Variance = 2

Recurrence relation for the Poisson distribution is

$$P(r+1) = \frac{\lambda}{r+1} P(r) = \frac{2}{r+1} P(r) \quad \dots(1)$$

$$\text{Now } P(r) = \frac{e^{-\lambda} \lambda^r}{r!} \Rightarrow P(0) = \frac{e^{-2} (2)^0}{0!} = e^{-2} = 0.1353$$

Putting $r = 0, 1, 2, 3$ in (1), we get

$$P(1) = 2P(0) = 2 \times 0.1353 = 0.2706; \quad P(2) = \frac{2}{2} P(1) = 0.2706$$

$$P(3) = \frac{2}{3} P(2) = \frac{2}{3} \times 0.2706 = 0.1804; \quad P(4) = \frac{2}{4} P(3) = \frac{1}{2} \times 0.1804 = 0.0902.$$

$$\text{Now, } P(r \geq 4) = 1 - [P(0) + P(1) + P(2) + P(3)]$$

$$= 1 - [0.1353 + 0.2706 + 0.2706 + 0.1804] = 0.1431.$$

Example 2. Using Poisson distribution, find the probability that the ace of spades will be drawn from a pack of well-shuffled cards at least once in 104 consecutive trials. (U.P.T.U. 2015)

$$\text{Sol. } p = \frac{1}{52}, n = 104$$

$$\therefore \lambda = np = 104 \times \frac{1}{52} = 2$$

$$\text{Prob. (at least once)} = P(r \geq 1) = 1 - P(0)$$

$$= 1 - \frac{e^{-\lambda} \cdot \lambda^0}{0!} = 1 - e^{-2} = 1 - 0.135335 \approx 0.8647.$$

Example 3. (i) Fit a Poisson distribution to the following data and calculate theoretical frequencies.

Deaths:	0	1	2	3	4
Frequencies:	122	260	15	2	1

[U.P.T.U. 2014 ; U.K.T.U. 2010]

(ii) The frequency of accidents per shift in a factory is shown in the following table:

Accident per shift	Frequency
0	192
1	100
2	24
3	3
4	1
Total	320

Calculate the mean number of accidents per shift. Fit a Poisson distribution and calculate theoretical frequencies.

$$\text{Sol. (i)} \text{ Mean of given distribution} = \frac{\sum fx}{\sum f}$$

$$\Rightarrow \lambda = \frac{60 + 30 + 6 + 4}{200} = 0.5$$

$$\text{Required Poisson distribution} = N \cdot \frac{e^{-\lambda} \cdot \lambda^r}{r!} = 200 \cdot \frac{e^{-0.5} (0.5)^r}{r!} = (121.306) \frac{(0.5)^r}{r!}$$

r	N. P(r)	Theoretical frequency
0	$121.306 \times \frac{(0.5)^0}{0!} = 121.306$	121
1	$121.306 \times \frac{(0.5)^1}{1!} = 60.653$	61
2	$121.306 \times \frac{(0.5)^2}{2!} = 15.163$	15
3	$121.306 \times \frac{(0.5)^3}{3!} = 2.527$	3
4	$121.306 \times \frac{(0.5)^4}{4!} = 0.3159$	0
		Total = 200

(ii) Mean number of accidents per shift

$$\lambda = \frac{\sum fx}{\sum f} = \frac{100 + 48 + 9 + 4}{320} = 0.5031$$

∴ Required Poisson distribution

$$= N \cdot \frac{e^{-\lambda} \cdot \lambda^r}{r!} = 320 \cdot \frac{e^{-.5031} (.5031)^r}{r!} = \frac{(193.48)(.5031)^r}{r!}$$

<i>r</i>	<i>N. P(r)</i>	<i>Theoretical frequency</i>
0	193.48	194
1	97.34	97
2	24.38	24
3	4.10	4
4	0.51	1
		Total = 320

Example 4. (i) Suppose that a book of 600 pages contains 40 printing mistakes. Assume that these errors are randomly distributed throughout the book and r , the number of errors per page has a Poisson distribution. What is the probability that 10 pages selected at random will be free from errors?

(ii) Wireless sets are manufactured with 25 solder joints each, on the average 1 joint in 500 is defective. How many sets can be expected to be free from defective joints in a consignment of 10000 sets?

Sol. (i) $p = \frac{40}{600} = \frac{1}{15}, n = 10$

$$\therefore \lambda = np = 10 \left(\frac{1}{15} \right) = \frac{2}{3}$$

$$P(r) = \frac{e^{-\lambda} \lambda^r}{r!} = \frac{e^{-2/3} (2/3)^r}{r!}$$

$$\therefore P(0) = \frac{e^{-2/3} (2/3)^0}{0!} = e^{-2/3} = 0.51.$$

(ii) $p = \frac{1}{500}, n = 25$

$$\therefore \lambda = np = 25 \times \frac{1}{500} = \frac{1}{20} = 0.05$$

No. of sets in a consignment, $N = 10000$

$$\text{Probability of having no defective joint} = P(r = 0) = \frac{e^{-0.05} (0.05)^0}{0!} = 0.9512.$$

∴ The expected no. of sets free from defective joints = $0.9512 \times 10000 = 9512$.

Example 5. A manufacturer knows that the condensors he makes contain on an average 1% of defectives. He packages them in boxes of 100. What is the probability that a box picked at random will contain 4 or more faulty condensors?

Sol. $p = 0.01, n = 100$

$$\therefore \lambda = np = 1$$

$$P(r) = \frac{e^{-\lambda} \lambda^r}{r!} = \frac{e^{-1}}{r!}$$

$$\begin{aligned}
 P(4 \text{ or more faulty condensors}) &= P(4) + P(5) + \dots + P(100) \\
 &= 1 - [P(0) + P(1) + P(2) + P(3)] \\
 &= 1 - \left[\frac{e^{-1}}{0!} + \frac{e^{-1}}{1!} + \frac{e^{-1}}{2!} + \frac{e^{-1}}{3!} \right] \\
 &= 1 - e^{-1} \left[1 + 1 + \frac{1}{2} + \frac{1}{6} \right] = 1 - \frac{8}{3e} = 1 - 0.981 = 0.019.
 \end{aligned}$$

Example 6. (i) If the probabilities of a bad reaction from a certain injection is 0.0002, determine the chance that out of 1000 individuals more than two will get a bad reaction.

(ii) The probability that a man aged 50 years will die within a year is 0.01125. What is the probability that of 12 such men, at least 11 will reach their 51st birthday?

(Given: $e^{-135} = 0.87366$)

Sol. (i) Here, $p = 0.0002$, $n = 1000$
 $\therefore \lambda = np = 1000 \times 0.0002 = 0.2$.

Since the prob. of bad reaction is very small and no. of trials is very high, we use Poisson distribution here.

The prob. that out of 100 individuals, more than 2 will get a bad reaction is

$$= P(r > 2) = 1 - P(r \leq 2) = 1 - [P(0) + P(1) + P(2)] \quad \dots(1)$$

Now, $P(0) = \frac{e^{-0.2} (0.2)^0}{0!} = 0.8187$ (Here $r = 0$)

$$P(1) = \frac{e^{-0.2} (0.2)^1}{1!} = 0.1637 \quad (\text{Here } r = 1)$$

and $P(2) = \frac{e^{-0.2} (0.2)^2}{2!} = 0.0164.$ (Here $r = 2$)

\therefore From (1), Reqd. probability $= 1 - [0.8187 + 0.1637 + 0.0164] = 0.0012.$

(ii) $p = 0.01125, n = 12$
 $\therefore \lambda = np = 12 \times 0.01125 = 0.135$

$P(\text{at least 11 survive}) = P(\text{atmost 1 dies})$

$$\begin{aligned}
 &= P(0) + P(1) = \frac{e^{-\lambda} \cdot \lambda^0}{0!} + \frac{e^{-\lambda} \cdot \lambda^1}{1!} \\
 &= e^{-0.135} (1 + 0.135) = 1.135 \times 0.87366 = 0.9916.
 \end{aligned}$$

Example 7. A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused ($e^{-1.5} = 0.2231$).

Sol. Since the number of demands for a car is distributed as a Poisson distribution with mean $\lambda = 1.5$.

\therefore Proportion of days on which neither car is used
 $=$ Probability of there being no demand for the car
 $= \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-1.5} = 0.2231$

Proportion of days on which some demand is refused
= probability for the number of demands to be more than two

$$= 1 - P(x \leq 2) = 1 - \left(e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} \right)$$

$$= 1 - e^{-1.5} \left(1 + 1.5 + \frac{(1.5)^2}{2} \right) = 0.1912625.$$

Example 8. Suppose the number of telephone calls on an operator received from 9 : 00 to 9 : 05 follow a Poisson distribution with a mean 3. Find the probability that

- (i) The operator will receive no calls in that time interval tomorrow.
 - (ii) In the next three days, the operator will receive a total of 1 call in that time interval.
- (Given: $e^{-3} = 0.04978$)

Sol. Here, $\lambda = 3$

$$(i) \quad P(0) = \frac{e^{-\lambda} \cdot \lambda^0}{0!} = e^{-3} = 0.04978$$

$$(ii) \quad \text{Reqd. probability} = P(0) P(0) P(1) + P(0) P(1) P(0) + P(1) P(0) P(0)$$

$$= 3 \left\{ \frac{e^{-\lambda} \cdot \lambda^0}{0!} \right\}^2 \frac{e^{-\lambda} \cdot \lambda^1}{1!} = 9(e^{-3})^3 = 0.00111.$$

Example 9. The no. of arrivals of customers during any day follows Poisson distribution with a mean of 5. What is the probability that the total no. of customers on two days selected at random is less than 2? (Given: $e^{-10} = 4.54 \times 10^{-5}$)

Sol. $\lambda = 5$

Arrival of Customers

I day	II day	Total
0	0	0
0	1	1
1	0	1

$$\begin{aligned} \text{Reqd. probability} &= P(0) P(0) + P(0) P(1) + P(1) P(0) \\ &= \frac{e^{-5} \cdot 5^0}{0!} \cdot \frac{e^{-5} \cdot 5^0}{0!} + \frac{e^{-5} \cdot 5^0}{0!} \cdot \frac{e^{-5} \cdot 5^1}{1!} + \frac{e^{-5} \cdot 5^1}{1!} \cdot \frac{e^{-5} \cdot 5^0}{0!} \\ &= e^{-10} + 2 \cdot 5 \cdot e^{-10} = 11 e^{-10} = 11 \times 4.54 \times 10^{-5} \\ &= 4.994 \times 10^{-4}. \end{aligned}$$

Example 10. An insurance company finds that 0.005% of the population dies from a certain kind of accident each year. What is the probability that the company must pay off no more than 3 of 10,000 insured risks against such incident in a given year?

$$\text{Sol.} \quad p = \frac{0.005}{100} = 0.00005, \quad n = 10000$$

$$\therefore \lambda = np = 10000 \times 0.00005 = 0.5$$

$$\text{Reqd. Probability} = 1 - P(r \leq 3) = 1 - [P(0) + P(1) + P(2) + P(3)]$$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{0!} + \frac{e^{-0.5}(0.5)^1}{1!} + \frac{e^{-0.5}(0.5)^2}{2!} + \frac{e^{-0.5}(0.5)^3}{3!} \right]$$

$$= 1 - e^{-5} [1 + 0.5 + 0.125 + 0.021] = 0.0016.$$

Example 11. In a certain factory turning out razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Calculate the approximate number of packets containing no defective, one defective and two defective blades in a consignment of 10,000 packets. (Given: $e^{-0.02} = 0.9802$) [U.P.T.U. 2009]

Sol. $p(\text{defective}) = 0.002$

$$n = 10 \quad (\text{no. of blades in a packet})$$

$$\therefore \lambda = np = 10 \times 0.002 = 0.02$$

$$\text{No. of packets in the consignment, } N = 10,000.$$

$$(i) \text{Probability of having no defective} = P(0) = \frac{e^{-0.02} (0.02)^0}{0!} = 0.9802 \quad | \text{ Here } r = 0$$

$$\text{Approximate no. of packets having zero defective in the consignment} = 0.9802 \times 10000 \\ = 9802$$

$$(ii) \text{Probability of having one defective} = P(1) = \frac{e^{-0.02} (0.02)^1}{1!} = 0.9802 \times 0.02 = 0.019604$$

$$\text{Approximate no. of packets having one defective in the consignment} \\ = 0.019604 \times 10000 \approx 196.$$

$$(iii) \text{Probability of having two defective blades}$$

$$P(2) = \frac{e^{-0.02} (0.02)^2}{2!} = \frac{(0.980198) \times (0.0004)}{2} = 0.000196.$$

$$\therefore \text{Approximate no. of packet having two defectives in the consignment} \\ = 0.000196 \times 10000 = 1.96 \approx 2.$$

Example 12. (i) Six coins are tossed 6400 times. Using the Poisson distribution, determine the approximate probability of getting six heads x times. [U.P.T.U. (C.O.) 2008]

(ii) A Poisson distribution has a double mode at $x = 3$ and $x = 4$. What is the probability that x will have one or the other of these two values? [U.P.T.U. (C.O.) 2008]

Sol. (i) Probability of getting one head with one coin = $\frac{1}{2}$.

$$\therefore \text{The probability of getting six heads with six coins} = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$$

$$\therefore \text{Average number of six heads with six coins in 6400 throws} = np = 6400 \times \frac{1}{64} = 100$$

$$\therefore \text{The mean of the Poisson distribution} = 100.$$

$$\text{Approximate probability of getting six heads } x \text{ times when the distribution is Poisson}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(100)^x \cdot e^{-100}}{x!}.$$

(ii) Since 2 modes are given when λ is an integer, modes are $\lambda - 1$ and λ .

$$\therefore \lambda - 1 = 3 \Rightarrow \lambda = 4$$

$$\text{Probability (when } r = 3) = \frac{e^{-4} (4)^3}{3!}$$

$$\text{Probability (when } r = 4) = \frac{e^{-4} (4)^4}{4!}$$

$$\therefore \text{Required Probability} = P(r = 3 \text{ or } 4) = P(r = 3) + P(r = 4)$$

$$= \frac{e^{-4} (4)^3}{3!} + \frac{e^{-4} (4)^4}{4!} = \frac{64}{3} e^{-4} = 0.39073.$$

Example 13. For a Poisson distribution with mean m , show that

$$\mu_{r+1} = mr \mu_{r-1} + m \frac{d\mu_r}{dm} \text{ where, } \mu_r = \sum_{x=0}^{\infty} (x-m)^r \frac{e^{-m} \cdot m^x}{x!}. \quad (\text{U.P.T.U. 2007})$$

$$\begin{aligned} \text{Sol. } \mu_r &= \sum_{x=0}^{\infty} (x-m)^r \cdot \frac{e^{-m} \cdot m^x}{x!} \\ \frac{d\mu_r}{dm} &= \sum_{x=0}^{\infty} \left[\frac{-e^{-m}}{x!} \cdot m^x (x-m)^r + \frac{e^{-m}}{x!} \{xm^{x-1} (x-m)^r - r(x-m)^{r-1} \cdot m^x\} \right] \\ \Rightarrow m \frac{d\mu_r}{dm} &= \sum_{x=0}^{\infty} \frac{e^{-m}}{x!} m^x (x-m)^{r+1} - rm \sum_{x=0}^{\infty} \frac{e^{-m}}{x!} m^x (x-m)^{r-1} = \mu_{r+1} - mr \mu_{r-1} \\ \Rightarrow \mu_{r+1} &= m \frac{d\mu_r}{dm} + mr \mu_{r-1}. \end{aligned}$$

Example 14. Show that in a Poisson distribution with unit mean, mean deviation about mean is $\left(\frac{2}{e}\right)$ times the standard deviation. (G.B.T.U. 2012)

Sol. Here, $\lambda = 1$

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

Mean deviation about mean 1 is

$$= \sum_{x=0}^{\infty} |x-1| p(x) = e^{-1} \sum_{x=0}^{\infty} \frac{|x-1|}{x!} = e^{-1} \left[1 + \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \dots \right] \quad \dots(1)$$

$$\text{we have, } \frac{n}{(n+1)!} = \frac{\overline{n+1}-1}{(n+1)!} = \frac{1}{n!} - \frac{1}{(n+1)!}$$

$$\begin{aligned} \therefore \text{From (1), Mean deviation about mean} &= e^{-1} \left[1 + \left(1 - \frac{1}{2!} \right) + \left(\frac{1}{2!} - \frac{1}{3!} \right) + \dots \right] \\ &= e^{-1} (1 + 1) = \frac{2}{3} \times 1 = \frac{2}{e} \times \text{S.D.} \quad | \text{ Since Variance} = \text{mean} = 1 \end{aligned}$$

ASSIGNMENT

1. If X is a Poisson variate such that $P(X = 2) = 9P(X = 4) + 90P(X = 6)$, find the standard deviation.
2. If a random variable has a Poisson distribution such that $P(1) = P(2)$, find
 - (i) mean of the distribution
 - (ii) $P(4)$.
3. Suppose that X has a Poisson distribution. If $P(X = 2) = \frac{2}{3} P(X = 1)$ find, (i) $P(X = 0)$ (ii) $P(X = 3)$.
4. A certain screw making machine produces on average 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.
5. (i) The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that in a group of 7, five or more will suffer from it?
 (ii) The experience shows that 4 industrial accidents occur in a plant on an average per month. Calculate the probabilities of less than 3 accidents in a certain month. Use Poisson distribution.
 (Given : $e^{-4} = 0.01832$). [M.T.U. (MBA) 2011]
6. (i) Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors ?
 (ii) Assume that the probability of an individual coalminer being killed in a mine accident during a year is $\frac{1}{2400}$. Use Poisson's distribution to calculate the probability that in a mine employing 200 miners there will be at least one fatal accident in a year.
7. (i) A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantee that not more than 10 pins will be defective, what is the approximate prob. that a box will fail to meet the guaranteed quality?
 (ii) An insurance company insures 4000 people against loss of both eyes in a car accident. Based on previous data it was assumed 10 persons out of 1,00,000 will have such type of injury in car accident. What is probability that more than 2 of the insured will collect on their policy in a given year? (M.T.U. 2013)
8. Records show that the probability is 0.00002 that a car will have a flat tyre while driving over a certain bridge. Use Poisson distribution to find the probability that among 20,000 cars driven over the bridge, not more than one will have a flat tyre.
9. Between the hours of 2 and 4 P.M., the average no. of phone calls per minute coming into the switch board of a company is 2.5. Find the probability that during a particular minute, there will be no phone call at all. [Given : $e^{-2} = 0.13534$ and $e^{-0.5} = 0.60650$.]
10. (i) Fit a Poisson distribution to the following data given the number of yeast cells per square for 400 squares :

No. of cells per sq. :	0	1	2	3	4	5	6	7	8	9	10
No. of squares :	103	143	98	42	8	4	2	0	0	0	0

 It is given that $e^{-1.3225} = 0.2665$.

 (ii) Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army corps. The distribution of deaths was as follows:

No. of deaths :	0	1	2	3	4	Total
Frequency :	109	65	22	3	1	200

 Fit a Poisson distribution to the data and calculate the theoretical frequencies. [M.T.U. (B. Pharma) 2011 ; M.T.U. (MBA) 2011]

 (iii) The following table gives the no. of days in a 50 day period during which automobile accidents occurred in a city.

No. of accidents :	0	1	2	3	4
No. of days :	21	18	7	3	1

 Fit a Poisson distribution to the data. (G.B.T.U. 2011)

Answers

3.61 NORMAL DISTRIBUTION

[U.P.T.U. 2007]

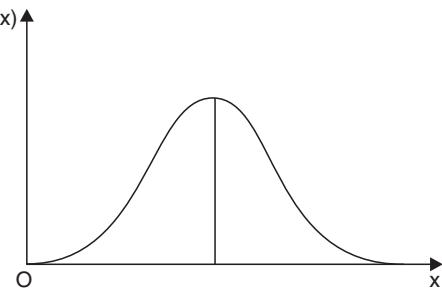
The normal distribution is a continuous distribution. It can be derived from the binomial distribution in the limiting case when n , the number of trials is very large and p , the probability of a success, is close to $\frac{1}{2}$. The general equation of the normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where the variable x can assume all values from $-\infty$ to $+\infty$. μ and σ , called the parameters of the distribution, are respectively the mean and the standard deviation of the distribution and $-\infty < \mu < \infty$, $\sigma > 0$. x is called the normal variate and $f(x)$ is called probability density function of the normal distribution.

If a variable x has the normal distribution with mean μ and standard deviation σ , we briefly write $x : N(\mu, \sigma^2)$.

The graph of the normal distribution is called the *normal curve*. It is bell-shaped and symmetrical about the mean μ . The two tails of the curve extend to $+\infty$ and $-\infty$ towards the positive and negative directions of the x -axis respectively and gradually approach the x -axis without ever meeting it. The curve is unimodal and the mode of the normal distribution coincides with its mean μ . The line $x = \mu$ divides the area under the normal curve above x -axis into two equal parts. Thus, the median of the distribution also coincides with its mean and mode. The area under the normal curve between any two given ordinates $x = x_1$ and $x = x_2$ represents the probability of values falling into the given interval. The total area under the normal curve above the x -axis is 1.



3.62 BASIC PROPERTIES OF THE NORMAL DISTRIBUTION

The probability density function of the normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

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

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

i.e., the total area under the normal curve above the x -axis is 1.

(iii) The normal distribution is symmetrical about its mean.

(iv) It is a unimodal distribution. The mean, mode and median of this distribution coincide.

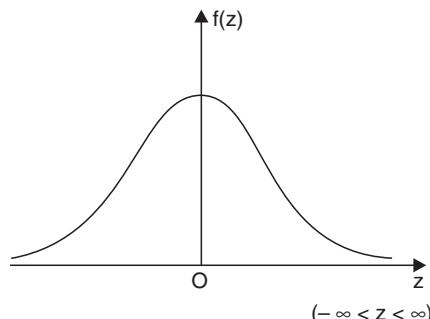
3.63 STANDARD FORM OF THE NORMAL DISTRIBUTION

If X is a normal random variable with mean μ and standard deviation σ , then the random variable $Z =$

$\frac{X - \mu}{\sigma}$ has the normal distribution with mean 0 and standard deviation 1. The random variable Z is called the *standardized (or standard) normal random variable*.

The probability density function for the normal distribution in standard form is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$



It is free from any parameter. This helps us to compute areas under the normal probability curve by making use of standard tables.

Note 1. If $f(z)$ is the probability density function for the normal distribution, then

$$P(z_1 \leq Z \leq z_2) = \int_{z_1}^{z_2} f(z) dz = F(z_2) - F(z_1), \quad \text{where } F(z) = \int_{-\infty}^z f(z) dz = P(Z \leq z)$$

The function $F(z)$ defined above is called the *distribution function* for the normal distribution.

Note 2. The probabilities $P(z_1 \leq Z \leq z_2)$, $P(z_1 < Z \leq z_2)$, $P(z_1 \leq Z < z_2)$ and $P(z_1 < Z < z_2)$ are all regarded to be the same.

Note 3. $F(-z_1) = 1 - F(z_1)$.

3.64 NORMAL DISTRIBUTION AS A LIMITING FORM OF BINOMIAL DISTRIBUTION (when $p = q$)

(U.P.T.U. 2007)

Let $N (q + p)^n$ be the binomial distribution. If $p = q$ then $p = q = \frac{1}{2}$ (since $p + q = 1$) and consequently the binomial distribution is symmetrical. Let n be an even integer say $2k$, k being an integer. Since $n \rightarrow \infty$, the frequencies of r and $r + 1$ successes can be written in following forms:

$$\begin{aligned} f(r) &= N \cdot {}^{2k}C_r \left(\frac{1}{2}\right)^{2k} \\ f(r+1) &= N \cdot {}^{2k}C_{r+1} \left(\frac{1}{2}\right)^{2k} \\ \therefore \frac{f(r+1)}{f(r)} &= \frac{{}^{2k}C_{r+1}}{{}^{2k}C_r} = \frac{2k-r}{r+1} \end{aligned}$$

The frequency of r successes will be greater than the frequency of $(r + 1)$ successes if

$$\begin{aligned} f(r) &> f(r+1) \\ \Rightarrow \frac{f(r+1)}{f(r)} &< 1 \\ \Rightarrow 2k-r &< r+1 \\ \Rightarrow r &> k - \frac{1}{2} \end{aligned} \quad \dots(1)$$

In a similar way, the frequency of r successes will be greater than the frequencies of $(r - 1)$ successes if $r < k + \frac{1}{2}$...(2)

In view of (1) and (2), we observe that if $k - \frac{1}{2} < r < k + \frac{1}{2}$ the frequency corresponding to r successes will be the greatest. Clearly, $r = k$ is the value of the success corresponding to which the frequency is maximum. Suppose it is y_0 . Then, we have

$$y_0 = N \cdot {}^{2k}C_k \left(\frac{1}{2}\right)^{2k} = N \cdot \frac{2k!}{k!k!} \left(\frac{1}{2}\right)^{2k}$$

Let y_x be the frequency of $k+x$ successes then, we have

$$y_x = N \cdot {}^{2k}C_{k+x} \left(\frac{1}{2}\right)^{2k} = N \cdot \left(\frac{1}{2}\right)^{2k} \cdot \frac{2k!}{(k+x)!(k-x)!}$$

Now,

$$\frac{y_x}{y_0} = \frac{k! k!}{(k+x)!(k-x)!} = \frac{k(k-1)(k-2)\dots(k-x+1)}{(k+x)(k+x-1)\dots(k+1)}$$

$$= \frac{\left(1 - \frac{1}{k}\right)\left(1 - \frac{2}{k}\right)\dots\left\{1 - \frac{x-1}{k}\right\}}{\left(1 + \frac{1}{k}\right)\left(1 + \frac{2}{k}\right)\dots\left(1 + \frac{x}{k}\right)}$$

Taking log on both sides,

$$\begin{aligned} \log \frac{y_x}{y_0} &= \left[\log\left(1 - \frac{1}{k}\right) + \log\left(1 - \frac{2}{k}\right) + \dots + \log\left(1 - \frac{x-1}{k}\right) \right] \\ &\quad - \left[\log\left(1 + \frac{1}{k}\right) + \log\left(1 + \frac{2}{k}\right) + \dots + \log\left(1 + \frac{x}{k}\right) \right] \quad \dots(3) \end{aligned}$$

Now, writing expression for each term and neglecting higher powers of $\frac{x}{k}$ (very small quantity), we get from (3),

$$\log \frac{y_x}{y_0} = -\frac{1}{k} \{1 + 2 + 3 + \dots + (x-1)\} - \frac{1}{k} \{1 + 2 + 3 + \dots + (x-1) + x\}$$

$$= -\frac{2}{k} \{1 + 2 + 3 + \dots + (x-1)\} - \frac{x}{k}$$

$$= -\frac{2}{k} \frac{(x-1)x}{2} - \frac{x}{k} = -\frac{x^2}{k}$$

$$\therefore y_x = y_0 e^{-x^2/k}$$

$$\Rightarrow \boxed{y_x = y_0 e^{-x^2/2\sigma^2}} \quad | \because \sigma^2 = npq = \frac{n}{4} = \frac{k}{2}$$

which is **normal distribution**.

3.65 MEAN AND VARIANCE OF NORMAL DISTRIBUTION

(U.P.T.U. 2015)

- The A.M. of a continuous distribution $f(x)$ is given by

$$\text{A.M. } (\bar{x}) = \frac{\int_{-\infty}^{\infty} x f(x) dx}{\int_{-\infty}^{\infty} f(x) dx} \quad | \text{ By definition}$$

Consider the normal distribution with μ, σ as the parameters then

$$\bar{x} = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad \left| \begin{array}{l} \text{Since } \int_{-\infty}^{\infty} f(x) dx \\ = \text{area under normal curve} = 1 \end{array} \right.$$

Put $\frac{x-\mu}{\sigma} = z$ so that $x = \mu + \sigma z \quad \therefore dx = \sigma dz$

so,

$$\begin{aligned} \bar{x} &= \int_{-\infty}^{\infty} (\mu + \sigma z) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} (\sigma dz) \\ &= \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz \\ &= \mu + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} d\left(\frac{z^2}{2}\right) \quad \left| \begin{array}{l} \therefore \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1 \end{array} \right. \\ &= \mu + \frac{\sigma}{\sqrt{2\pi}} \left(\frac{e^{-z^2/2}}{-1} \right)_{-\infty}^{\infty} \\ &= \boxed{\bar{x} = \mu} \end{aligned}$$

2. By definition,

$$\begin{aligned} \text{Variance} &= \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx + \bar{x}^2 \int_{-\infty}^{\infty} f(x) dx - 2\bar{x} \int_{-\infty}^{\infty} xf(x) dx \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx + \bar{x}^2 - 2\bar{x}\bar{x} \quad \left| \begin{array}{l} \therefore \int_{-\infty}^{\infty} f(x) dx = 1 \text{ and} \\ \int_{-\infty}^{\infty} xf(x) dx = \bar{x} \end{array} \right. \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - \bar{x}^2 \end{aligned} \quad \dots(1)$$

Now, Let $I = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} dx$

Put $\frac{x-\bar{x}}{\sigma} = z$ so that $x = \bar{x} + \sigma z \quad \therefore dx = \sigma dz$

$$\begin{aligned} \text{Hence, } I &= \int_{-\infty}^{\infty} (\bar{x} + \sigma z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} \sigma dz \\ &= \frac{1}{\sqrt{2\pi}} \left[\sigma^2 \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz + \bar{x}^2 \int_{-\infty}^{\infty} e^{-z^2/2} dz + 2\sigma\bar{x} \int_{-\infty}^{\infty} z e^{-z^2/2} dz \right] \end{aligned}$$

$$\begin{aligned}
 &= \frac{-\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z d(e^{-z^2/2}) + \bar{x}^2 \cdot 1 + 2\sigma\bar{x} \cdot 0 \\
 &= -\frac{\sigma^2}{\sqrt{2\pi}} \left(ze^{-z^2/2}\right)_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz + \bar{x}^2 \\
 &= 0 + \sigma^2 \cdot 1 + \bar{x}^2 = \sigma^2 + \bar{x}^2
 \end{aligned}$$

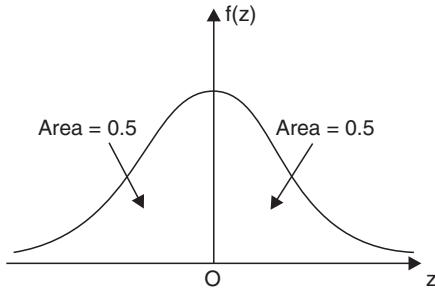
∴ From (1), Variance = $\sigma^2 + \bar{x}^2 - \bar{x}^2 = \sigma^2$

∴ The standard deviation of the normal distribution is σ .

3.66 AREA UNDER THE NORMAL CURVE

By taking $z = \frac{x - \mu}{\sigma}$, standard normal curve is formed.
The total area under this curve is 1.

The area under the curve is divided into two equal parts by $z = 0$. The area between the ordinate $z = 0$ and any other ordinate can be noted from the supplied table. It should be noted that mean for the normal distribution is 0.



3.67 APPLICATIONS OF NORMAL DISTRIBUTION

De Moivre made the discovery of this distribution in 1733.

This distribution has an important application in the theory of errors made by chance in experimental measurements. Its more applications are in computation of hit probability of a shot and in statistical inference in almost every branch of science.

EXAMPLES

Example 1. A sample of 100 dry battery cells tested to find the length of life produced the following results:

$$\bar{x} = 12 \text{ hours}, \sigma = 3 \text{ hours.}$$

Assuming the data to be normally distributed, what percentage of battery cells are expected to have life

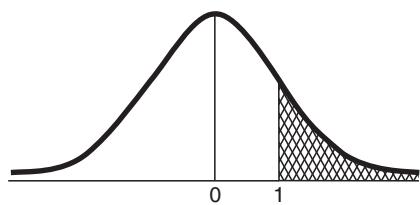
- (i) more than 15 hours (ii) less than 6 hours (iii) between 10 and 14 hours?

Sol. Here x denotes the length of life of dry battery cells.

$$\text{Also } z = \frac{x - \bar{x}}{\sigma} = \frac{x - 12}{3}.$$

(i) When $x = 15$, $z = 1$

$$\begin{aligned}
 \therefore P(x > 15) &= P(z > 1) \\
 &= P(0 < z < \infty) - P(0 < z < 1) \\
 &= .5 - 0.3413 = 0.1587 = 15.87\%.
 \end{aligned}$$



(ii) When $x = 6$, $z = -2$

$$\begin{aligned}\therefore P(x < 6) &= P(z < -2) \\ &= P(z > 2) = P(0 < z < \infty) - P(0 < z < 2) \\ &= 0.5 - 0.4772 = 0.0228 = 2.28\%.\end{aligned}$$

(iii) When $x = 10$, $z = -\frac{2}{3} = -0.67$

When $x = 14$, $z = \frac{2}{3} = 0.67$

$$\begin{aligned}P(10 < x < 14) &= P(-0.67 < z < 0.67) \\ &= 2P(0 < z < 0.67) = 2 \times 0.2485 \\ &= 0.4970 = 49.70\%.\end{aligned}$$

Example 2. In a sample of 1000 cases, the mean of a certain test is 14 and S.D. is 2.5. Assuming the distribution to be normal, find

(i) how many students score between 12 and 15?

(ii) how many score above 18?

(iii) how many score below 8?

(iv) how many score 16?

Sol. (i)

$$\begin{aligned}z_1 &= \frac{x_1 - \mu}{\sigma} = \frac{12 - 14}{2.5} = -0.8 \\ z_2 &= \frac{x_2 - \mu}{\sigma} = \frac{15 - 14}{2.5} = 0.4\end{aligned}$$

Area lying between -0.8 and 0.4

$$\begin{aligned}&= \text{Area between } 0 \text{ to } 0.8 + \text{Area between } 0 \text{ to } 0.4 \\ &= 0.2881 + 0.1554 = 0.4435\end{aligned}$$

$$\text{Reqd. no. of students} = 1000 \times 0.4435 = 444 \text{ (app.)}$$

(ii)

$$z = \frac{18 - 14}{2.5} = 1.6$$

$$\text{Area right to } 1.6 = 0.5 - (\text{Area between } 0 \text{ and } 1.6) = 0.5 - 0.4452 = 0.0548$$

$$\text{Reqd. no. of students} = 1000 \times 0.0548 = 54.8 \approx 55 \text{ (app.)}$$

(iii)

$$z = \frac{8 - 14}{2.5} = -2.4$$

$$\text{Area left to } -2.4 = 0.5 - (\text{Area between } 0 \text{ and } 2.4) = 0.5 - 0.4918 = 0.0082$$

$$\therefore \text{Reqd. no. of students} = 1000 \times 0.0082 = 8.2 \approx 8 \text{ (app.)}$$

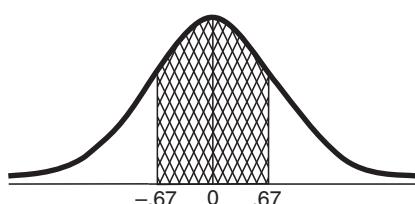
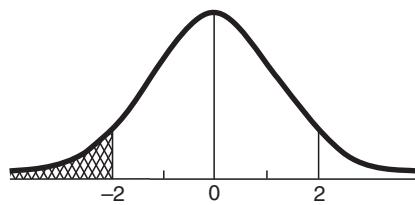
(iv)

$$z_1 = \frac{15.5 - 14}{2.5} = 0.6$$

$$z_2 = \frac{16.5 - 14}{2.5} = 1$$

$$\text{Area between } 0.6 \text{ and } 1 = 0.3413 - 0.2257 = 0.1156$$

$$\therefore \text{Reqd. no. of students} = 1000 \times 0.1156 = 115.6 \approx 116 \text{ (app.)}.$$



Example 3. Assume mean height of soldiers to be 68.22 inches with a variance of 10.8 inches square. How many soldiers in a regiment of 1,000 would you expect to be over 6 feet tall, given that the area under the standard normal curve between $z = 0$ and $z = 0.35$ is 0.1368 and between $z = 0$ and $z = 1.15$ is 0.3746.
[G.B.T.U. (C.O.) 2011]

Sol. $x = 6$ feet = 72 inches

$$\therefore z = \frac{x - \mu}{\sigma} = \frac{72 - 68.22}{\sqrt{10.8}} = 1.15$$

$$\begin{aligned} P(x > 72) &= P(z > 1.15) = 0.5 - P(0 \leq z \leq 1.15) \\ &= 0.5 - 0.3746 = 0.1254 \end{aligned}$$

\therefore Expected no. of soldiers = $1000 \times 0.1254 = 125.4 \approx 125$ (app.).

Example 4. A large number of measurement is normally distributed with a mean 65.5" and S.D. of 6.2". Find the percentage of measurements that fall between 54.8" and 68.8".

Sol. Mean $\mu = 65.5$ inches, S.D. $\sigma = 6.2$ inches

$$x_1 = 54.8 \text{ inches}, x_2 = 68.8 \text{ inches}$$

$$\therefore z_1 = \frac{x_1 - \mu}{\sigma} = \frac{54.8 - 65.5}{6.2} = -1.73$$

$$\text{and } z_2 = \frac{x_2 - \mu}{\sigma} = \frac{68.8 - 65.5}{6.2} = 0.53$$

$$\begin{aligned} \text{Now, } P(-1.73 \leq z \leq 0.53) &= P(-1.73 \leq z \leq 0) + P(0 \leq z \leq 0.53) \\ &= P(0 \leq z \leq 1.73) + P(0 \leq z \leq 0.53) \\ &= 0.4582 + 0.2019 = 0.6601 \end{aligned}$$

| By table

\therefore Reqd. percentage of measurements = 66.01%.

Example 5. A manufacturer knows from experience that the resistance of resistors he produces is normal with mean $\mu = 100$ ohms and standard deviation $\sigma = 2$ ohms. What percentage of resistors will have resistance between 98 ohms and 102 ohms?

Sol. $\mu = 100 \Omega, \sigma = 2 \Omega, x_1 = 98 \Omega, x_2 = 102 \Omega$

$$\therefore z_1 = \frac{x_1 - \mu}{\sigma} = \frac{98 - 100}{2} = -1$$

$$\text{and } z_2 = \frac{x_2 - \mu}{\sigma} = \frac{102 - 100}{2} = 1.$$

$$\begin{aligned} \text{Now, } P(98 < x < 102) &= P(-1 < z < 1) \\ &= P(-1 \leq z \leq 0) + P(0 \leq z \leq 1) \\ &= P(0 \leq z \leq 1) + P(0 \leq z \leq 1) \\ &= 0.3413 + 0.3413 = 0.6826. \end{aligned}$$

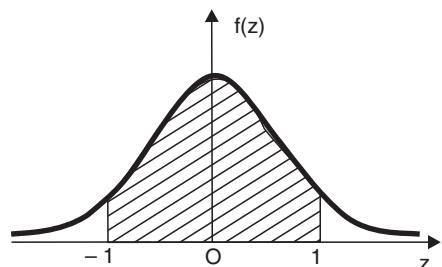
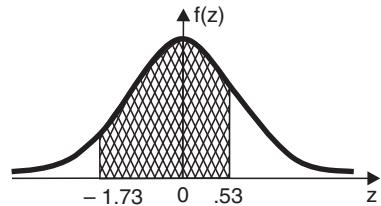
\therefore Percentage of resistors having resistance between 98Ω and $102 \Omega = 68.26\%$.

Example 6. In a normal distribution, 31% of the items are under 45 and 8% are over 64.

Find the mean and standard deviation of the distribution. It is given that if $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{1}{2}x^2} dx$ then $f(0.5) = 0.19$ and $f(1.4) = 0.42$.
(M.T.U. 2013)

Sol. Let μ and σ be the mean and S.D. respectively.

31% of the items are under 45.



\Rightarrow Area to the left of the ordinate $x = 45$ is 0.31

When $x = 45$, let $z = z_1$

$$P(z_1 < z < 0) = 0.5 - 0.31 = 0.19$$

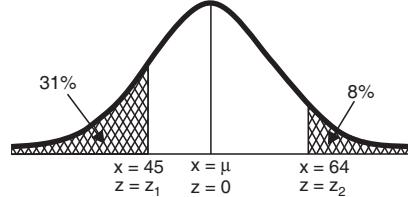
From the tables, the value of z corresponding to this area is 0.5

$$\therefore z_1 = -0.5 [z_1 < 0]$$

When $x = 64$, let $z = z_2$

$$P(0 < z < z_2) = 0.5 - 0.08 = 0.42$$

From the tables, the value of z corresponding to this area is 1.4.



$$\therefore z_2 = 1.4$$

Since

$$z = \frac{x - \mu}{\sigma}$$

$$\therefore -0.5 = \frac{45 - \mu}{\sigma} \quad \text{and} \quad 1.4 = \frac{64 - \mu}{\sigma}$$

$$\Rightarrow 45 - \mu = -0.5\sigma \quad \dots(1)$$

$$\text{and} \quad 64 - \mu = 1.4\sigma \quad \dots(2)$$

$$\text{Subtracting} \quad -19 = -1.9\sigma \quad \therefore \sigma = 10$$

$$\text{From (1),} \quad 45 - \mu = -0.5 \times 10 = -5 \quad \therefore \mu = 50.$$

Example 7. The life of army shoes is normally distributed with mean 8 months and standard deviation 2 months. If 5000 pairs are insured, how many pairs would be expected to need replacement after 12 months? $\left[\text{Given that } P(z \geq 2) = 0.0228 \text{ and } z = \frac{x - \mu}{\sigma} \right]$.

Sol. Mean (μ) = 8, Standard Deviation (σ) = 2

Number of pairs of shoes = 5000, Total months (x) = 12

$$\text{when } x = 12, \quad z = \frac{x - \mu}{\sigma} = \frac{12 - 8}{2} = 2$$

$$\text{Area } (z \geq 2) = 0.0228$$

$$\text{Number of pairs whose life is more than 12 months} = 5000 \times 0.0228 = 114$$

$$\text{Pair of shoes needing replacement after 12 months} = 5000 - 114 = 4886.$$

Example 8. The mean inside diameter of a sample of 200 washers produced by a machine is 0.502 cm and the standard deviation is 0.005 cm. The purpose for which these washers are intended allows a minimum tolerance in the diameter of 0.496 to 0.508 cm, otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine. Assume the diameters are normally distributed.

Sol. Given: Mean $\mu = 0.502$ cm, S.D. $\sigma = 0.005$ cm, $x_1 = 0.496$ cm, $x_2 = 0.508$ cm.

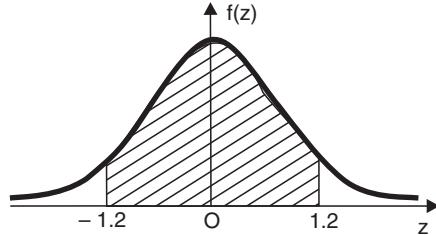
$$\text{Now, } z_1 = \frac{x_1 - \mu}{\sigma} = \frac{0.496 - 0.502}{0.005} = -1.2$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{0.508 - 0.502}{0.005} = 1.2$$

Area for non-defective washers

$$\begin{aligned}
 &= P(-1.2 \leq z \leq 1.2) \\
 &= P(-1.2 \leq z \leq 0) + P(0 \leq z \leq 1.2) \\
 &= P(0 \leq z \leq 1.2) + P(0 \leq z \leq 1.2) \\
 &= 0.3849 + 0.3849 = 0.7698 \\
 &= 76.98\%.
 \end{aligned}$$

$$\begin{aligned}
 \therefore \text{Percentage of defective washers} &= 100 - 76.98 \\
 &= 23.02\%.
 \end{aligned}$$



Example 9. Assuming that the diameters of 1000 brass plugs taken consecutively from a machine, form a normal distribution with mean 0.7515 cm and standard deviation 0.002 cm, how many of the plugs are likely to be rejected if the approved diameter is 0.752 ± 0.004 cm.

Sol. Tolerance limits of the diameter of non-defective plugs are

$$0.752 - 0.004 = 0.748 \text{ cm. and } 0.752 + 0.004 = 0.756 \text{ cm.}$$

$$\text{Standard normal variable, } z = \frac{x - \mu}{\sigma}$$

$$\text{If } x_1 = 0.748, \quad z_1 = \frac{0.748 - 0.7515}{0.002} = -1.75$$

$$\text{If } x_2 = 0.756, \quad z_2 = \frac{0.756 - 0.7515}{0.02} = 2.25$$

Area from $(z_1 = -1.75)$ to $(z_2 = 2.25)$

$$\begin{aligned}
 &= P(-1.75 \leq z \leq 2.25) = P(-1.75 \leq z \leq 0) + P(0 \leq z \leq 2.25) \\
 &= P(0 \leq z \leq 1.75) + P(0 \leq z \leq 2.25) = 0.4599 + 0.4878 = 0.9477
 \end{aligned}$$

Number of plugs which are likely to be rejected = $1000 \times (1 - 0.9477) = 1000 \times 0.0523 = 52.3$

Hence approximately 52 plugs are likely to be rejected.

Example 10. If the heights of 300 students are normally distributed with mean 64.5 inches and standard deviation 3.3 inches, find the height below which 99% of the students lie.

Sol. Mean $\mu = 64.5$ inches, S.D. $\sigma = 3.3$ inches

$$\text{Area between 0 and } \frac{x - 64.5}{3.3} = 0.99 - 0.5 = 0.49$$

From the table, for the area 0.49, $z = 2.327$

The corresponding value of x is given by

$$\begin{aligned}
 \frac{x - 64.5}{3.3} &= 2.327 \\
 \Rightarrow x - 64.5 &= 7.68 \\
 \Rightarrow x &= 7.68 + 64.5 = 72.18 \text{ inches.}
 \end{aligned}$$

Hence 99% students are of height less than 6 ft. 0.18 inches.

Example 11. The income of a group of 10,000 persons was found to be normally distributed with mean ₹ 750 p.m. and standard deviation of ₹ 50. Show that, of this group, about 95% had income exceeding ₹ 668 and only 5% had income exceeding ₹ 832. Also find the lowest income among the richest 100.

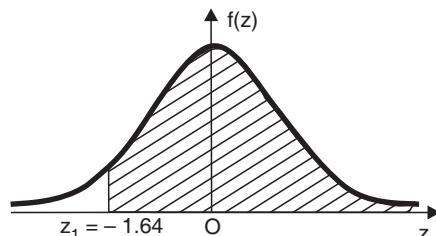
Sol. Given: $\mu = 750, \sigma = 50$

Standard normal variable, $z = \frac{x - \mu}{\sigma}$

$$(i) \text{ If } x_1 = 668, z_1 = \frac{x_1 - \mu}{\sigma} = \frac{668 - 750}{50} = -1.64$$

$$\begin{aligned} P(x_1 > 668) &= P(z_1 > -1.64) \\ &= 0.5 + P(-1.64 \leq z \leq 0) \\ &= 0.5 + P(0 \leq z \leq 1.64) \\ &= 0.5 + 0.4495 \\ &= 0.9495 \end{aligned}$$

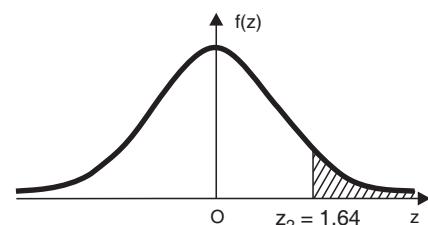
\therefore Required percentage of persons having income exceeding ₹ 668 = 94.95% \approx 95% (approx.)



$$(ii) \text{ If } x_2 = 832, z_2 = \frac{x_2 - \mu}{\sigma} = \frac{832 - 750}{50} = 1.64$$

$$\begin{aligned} P(x_2 > 832) &= P(z_2 > 1.64) \\ &= 0.5 - P(0 \leq z \leq 1.64) \\ &= 0.5 - 0.4495 = 0.0505 \end{aligned}$$

\therefore Required percentage of persons having income exceeding ₹ 832 = 5.05% \approx 5% (approx.)



(iii) Let x be the lowest income among the richest 100 persons i.e., 1% of 10,000.

Thus, area between O and $z = 0.49$ (see figure) by Normal distribution table,

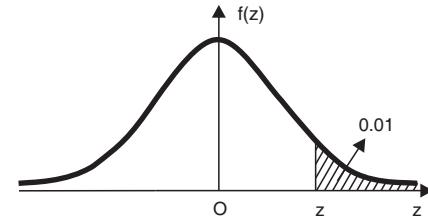
$$z = 2.33$$

Thus,

$$\frac{x - \mu}{\sigma} = 2.33$$

$$\Rightarrow \frac{x - 750}{50} = 2.33$$

$$\Rightarrow x = 866.5$$



Hence ₹ 866.5 is the minimum income among the richest 100 persons.

Example 12. 255 metal rods were cut roughly 6 inches over size. Finally the lengths of the over size amount, were measured exactly and grouped with 1 inch intervals, there being in

all 12 groups $\frac{1}{2}'' - 1\frac{1}{2}'', 1\frac{1}{2}'' - 2\frac{1}{2}'', \dots, 11\frac{1}{2}'' - 12\frac{1}{2}''$.

The frequency distribution for the 255 lengths was as follows:

Length (inches)	1	2	3	4	5	6	7	8	9	10	11	12
Central value												
Frequency	2	10	19	25	40	44	41	28	25	15	5	1

Fit a normal curve to this data.

Sol. The equation of the normal curve for N observations is

$$y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \dots(1)$$

x	f	$u = x - 6$	fu	fu^2
1	2	-5	-10	50
2	10	-4	-40	160
3	19	-3	-57	171
4	25	-2	-50	100
5	40	-1	-40	40
6	44	0	0	0
7	41	1	41	41
8	28	2	56	112
9	25	3	75	225
10	15	4	60	240
11	5	5	25	125
12	1	6	6	36
Total	255		66	1300

Mean, $\mu = a + \frac{\Sigma fu}{\Sigma f} = 6 + \frac{66}{255} = 6.259$

Variance, $\sigma^2 = \frac{\Sigma fu^2}{\Sigma f} - \left(\frac{\Sigma fu}{\Sigma f} \right)^2 = \frac{1300}{225} - \left(\frac{66}{255} \right)^2 = 5.031$

$\therefore \sigma = 2.243$

Thus, we have $N = 255$, Mean, $\mu = 6.259$ ", S.D. $\sigma = 2.243$ "

Hence the fitted curve is

$$y = \frac{255}{2.243\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-6.259}{2.243}\right)^2} \quad | \text{ From (1)}$$

$$= \frac{113.68}{\sqrt{2\pi}} e^{-0.099(x-6.259)^2}$$

Example 13. Show that the area under the normal curve is unity.

Sol. Area under the normal curve is given by

$$A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Put $\frac{x-\mu}{\sigma} = z$ so $dx = \sigma dz$

$$\therefore A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} (\sigma dz) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-z^2/2} dz$$

Now,

$$A \cdot A = A^2 = \left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-x^2/2} dx \right) \left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-y^2/2} dy \right)$$

$$= \frac{2}{\pi} \int_0^{\infty} \int_0^{\infty} e^{-\frac{(x^2+y^2)}{2}} dx dy \quad | \text{ where } x \text{ and } y \text{ are dummy variables}$$

Put $x = r \cos \theta$, $y = r \sin \theta$ so that $J = r$ changing to polar coordinates,

$$A^2 = \frac{2}{\pi} \int_0^{\pi/2} \int_0^\infty e^{-r^2/2} r dr d\theta = \int_0^\infty e^{-r^2/2} d\left(\frac{r^2}{2}\right) = 1$$

$\therefore A = \text{Area under the normal curve} = 1$

Example 14. Prove that for normal distribution, the mean deviation from the mean equals to $\frac{4}{5}$ of the standard deviation approximately. (U.P.T.U. 2009)

Sol. Let μ and σ be the mean and standard deviation of the normal distribution. Then by definition,

Mean deviation from the mean

$$\begin{aligned} &= \int_{-\infty}^{\infty} |x - \mu| f(x) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |x - \mu| e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma |z| e^{-\frac{1}{2}z^2} \sigma dz \quad \left| \begin{array}{l} \text{when } \frac{x-\mu}{\sigma} = z \\ \Rightarrow dx = \sigma dz \end{array} \right. \\ &= \sigma \sqrt{\frac{2}{\pi}} \int_0^{\infty} z e^{-z^2/2} dz \\ &= \sigma \sqrt{\frac{2}{\pi}} \left[-e^{-z^2/2} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}} \sigma = 0.7979 \sigma \approx 0.8\sigma \approx \frac{4}{5}\sigma \end{aligned}$$

ASSIGNMENT

1. In a test on 2000 electric bulbs, it was found that the life of a particular make, was normally distributed with an average life of 2040 hours and S.D. of 60 hours, estimate the number of bulbs likely to burn for
 - (i) more than 2150 hours
 - (ii) less than 1950 hours
 - (iii) more than 1920 hours but less than 2160 hours. (U.P.T.U. 2008)
2. An aptitude test for selecting officers in a bank is conducted on 1000 candidates. The average score is 42 and the standard deviation of score is 24. Assuming normal distribution for the scores, find
 - (i) the number of candidates whose scores exceed 60
 - (ii) the number of candidates whose scores lie between 30 and 60.
3. (i) In a normal distribution exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution? (G.B.T.U. 2010)

 (ii) In a normal distribution, 0.0107 of the items lie below 42 and 0.0446 of the items lie above 82. What is the mean and standard deviation of the normal distribution?

[U.P.T.U. (MBA) 2009]
4. If Z is a standard normal variable, find the following probabilities: [G.B.T.U. (MBA) 2010]
 - (i) $P(Z < 1.2)$
 - (ii) $P(Z > -1.2)$
 - (iii) $P(-1.2 < Z < 1.3)$.

5. An aptitude test was conducted on 900 employees of the Metro Tyres Limited in which the mean score was found to be 50 units and standard deviation was 20. On the basis of this information, you are required to answer the following questions:

- (i) What was the number of employees whose mean score was less than 30?
- (ii) What was the number of employees whose mean score exceeded 70?
- (iii) What was the number of employees whose mean score were between 30 and 70?

$\frac{x - \mu}{\sigma}$	0.25	0.50	0.70	1.00	1.25	1.50
Area	0.0987	0.1915	0.2734	0.3413	0.3944	0.4332

[U.P.T.U. (MBA) 2009]

6. (a) Students of a class were given a mechanical aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percent of students scored?

- (i) more than 60 marks? (ii) less than 56 marks? (iii) between 45 and 65 marks?

- (b) 2000 students appeared in an examination. Distribution of marks is assumed to be normal with mean $\mu = 30$ and $\sigma = 6.25$. How many students are expected to get marks?

- (i) between 20 and 40 (ii) less than 35 and (iii) above 50.

[U.P.T.U. (MBA) 2012]

- (c) Suppose the weight W of 600 male students are normally distributed with mean $\mu = 70$ kg and standard deviation $\sigma = 5$ kg. Find number of students with weight

- (i) between 69 and 74 kg (ii) more than 76 kg.

(G.B.T.U. 2013)

7. (a) In an intelligence test administered to 1000 students, the average score was 42 and standard deviation 24. Find:

- (i) the expected number of students scoring more than 50.

- (ii) the number of students scoring between 30 and 54.

- (iii) the value of score exceeded by top 100 students. [G.B.T.U. (MBA) 2010]

- (b) The average monthly sales of 5000 firms are normally distributed. Its mean and standard deviation are ₹ 36000 and ₹ 10000 respectively. Find:

- (i) the no. of firms having sales over ₹ 40000.

- (ii) the no. of firms having sales between ₹ 30000 and ₹ 40000.

[Given area under normal curve from 0 to z for $Z(0.4) = 0.1554$ and $Z(0.6) = 0.2257$]

[G.B.T.U. (MBA) 2010]

- (c) The daily wages of 1000 workers are distributed around a mean of ₹ 140 and with a standard deviation of ₹ 10. Estimate the number of workers whose daily wages will be

- (i) between ₹ 140 and ₹ 144 (ii) less than ₹ 126

- (iii) more than ₹ 160. (G.B.T.U. 2012)

8. (a) Records kept by the goods inwards department of a large factory show that the average no. of lorries arriving each week is 248. It is known that the distribution approximates to be normal with a standard deviation of 26.

If this pattern of arrival continues, what percentage of weeks can be expected to have number of arrivals of:

- (i) less than 229 per week? (ii) more than 280 per week?

- (b) Pipes for tobacco are being packed in fancy plastic boxes. The length of the pipe is normally distributed with $\mu = 5"$ and $\sigma = 0.1"$. The internal length of the boxes is 5.2". What is the probability that the box would be small for the pipe?

[Given that : $\phi(1.8) = 0.9641$, $\phi(2) = 0.9772$, $\phi(2.5) = 0.9938$]

- (c) A manufacturer of envelopes knows that the weight of the envelopes is normally distributed with mean 1.9 gm and variance 0.01 square gm. Find how many envelopes weighing

- (i) 2 gm or more

17. How does a normal distribution differ from a binomial distribution? What are the important properties of normal distribution? [M.T.U. (MBA) 2012]
18. If the skulls are classified as A, B and C according as the length-breadth index is under 75, between 75 and 80 or over 80, find approximately (assuming that the distribution is normal) the mean and standard deviation of a series in which A are 58%, B are 38% and C are 4%, being given

that if $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-(x^2/2)} dx$ then $f(0.20) = 0.08$ and $f(1.75) = 0.46$.

[Hint: $P(X < 75) = 0.58$, $P(X > 80) = 0.04$]

19. The following table gives frequencies of occurrence of a variable X between certain limits:

Variable X	Frequency
Less than 40	30
40 or more but less than 50	33
50 and more	37

The distribution is exactly normal. Find the distribution and also obtain the frequency between $X = 50$ and $X = 60$.

20. The marks X obtained in Mathematics by 1000 students are normally distributed with mean 78% and standard deviation 11%.

Determine:

- (i) how many students got marks above 90%?
- (ii) What was the highest marks obtained by the lowest 10% of students?
- (iii) Within what limits did the middle 90% of the students lie?

Answers

- | | |
|--------------------------------------------------|-------------------------------------------|
| 1. (i) 67 (ii) 134 (iii) 1909 | 2. (i) 227 (ii) 465 |
| 3. (i) $\bar{x} = 50.3$, $\sigma = 10.33$ | (ii) $\mu = 65$, $\sigma = 10$ |
| 4. (i) 0.8849 | (ii) 0.8849 |
| 5. (i) 143 | (ii) 143 |
| 6. (a) (i) 50%, (ii) 21.2%, (iii) 84% | (b) (i) 1781, (ii) 1576, (iii) 1 |
| 7. (a) (i) 371, (ii) 383, (iii) 72.72 | (b) (i) 1723 (ii) 1906 |
| 8. (a) (i) 23% (ii) 11% | (b) 0.0228 |
| 9. (a) (i) 48 (ii) 251 (iii) 701 | (b) 294 |
| 10. (i) 79 (ii) 35% (iii) 11 | 11. 10,000 |
| 13. 0.06357 | 14. 34% |
| 16. (i) 3.85 (ii) 47.4. | 18. $\mu = 74.35$, $\sigma = 3.23$ |
| 20. (i) 138 (ii) 63.92% (iii) between 60 and 96. | 19. $\mu = 46.12$, $\sigma = 11.76$, 25 |

3.68 POPULATION OR UNIVERSE

An aggregate of objects (animate or inanimate) under study is called **population or universe**. It is thus a collection of individuals or of their attributes (qualities) or of results of operations which can be numerically specified.

A universe containing a finite number of individuals or members is called a **finite universe**. For example, the universe of the weights of students in a particular class.

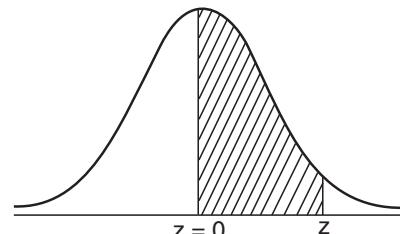
A universe with infinite number of members is known as an **infinite universe**. For example, the universe of pressures at various points in the atmosphere.

In some cases, we may be even ignorant whether or not a particular universe is infinite, e.g., the universe of stars.

APPENDIX

**Table 1: NORMAL TABLE
AREAS UNDER THE STANDARD NORMAL**

$$\text{CURVE} = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{z^2}{2}} dz$$



<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2485	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4255	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4930	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4999	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

(i)

**Table 2 : SIGNIFICANT VALUES $t_v(\alpha)$ OF t-DISTRIBUTION
(TWO TAIL AREAS) [$|t| > t_v(\alpha)$] = α**

d.f. (v)	Probability (Level of Significance)					
	0.50	0.10	0.05	0.02	0.01	0.001
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

Table 3 : F-Distribution
Values of F for F-Distributions with 0.05 of the Area in the Right Tail

	<i>Degrees of freedom for numerator</i>																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	9.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	3.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	3.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.17	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.98	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.94	2.40	2.34	2.28	2.24	2.16	2.29	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.64	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Table 4 : CHI-SQUARE
Significant Values $\chi^2(\alpha)$ of Chi-Square Distribution Right Tail Areas
for Given Probability α ,
 $P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$
And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	Probability (Level of Significance)						
	0.99	0.95	0.50	0.10	0.05	0.02	0.01
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Note. For degrees of freedom (v) greater than 30, the quantity $\sqrt{2\chi^2} - \sqrt{2v-1}$ may be used as a normal variate with unit variance.