CHAPTER 26

Probability and Distributions

26.1 (1) INTRODUCTION

We often hear such statements : 'It is likely to rain today', 'I have a fair chance of getting admission', and 'There is an even chance that in tossing a coin the head may come up'. In each case, we are not certain of the outcome, but we wish to assess the chances of our predictions coming true. The study of probability provides a mathematical framework for such assertions and is essential in every decision making process. Before defining probability, let us explain a few terms :

(2) Principle of counting. If an event can happen in n_1 ways and thereafter for each of these events a second event can happen in n_2 ways, and for each of these first and second events a third event can happen for n_3 ways and so on, then the number of ways these m event can happen is given by the product $n_1 . n_2 . n_3 ... n_m$.

(3) Permutations. A permutation of a number of objects is their arrangement in some definite order. Given three letters a, b, c, we can permute them two at a time as "bc, cb ; ca, ac; ab, ba" yielding 6 permutations. The combinations or groupings are only 3, *i.e.*, bc, ca, ab. Here the order is immetrial.

The number of permutations of n different thing taken r at a time is

$$n(n-1)(n-2)\ldots(n-r+1), \text{ which is denoted by } {}^nP_r.$$

Thus
$$ {}^nP_r = n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!}$$

Permutations with repetitions. The number of permutations of n objects of which n_1 are alike, n_2 are alike and n_3 are alike is $\dfrac{n!}{n_1!\,n_2!\,n_3!}$.

(4) Combinations. *The number of combinations of n different objects taken r at a time is denoted by* ${}^nC_r$. If we take any one of the combinations, its r objects can be arranged in $r!$ ways. So the total number of arrangements which can be obtained from all the combinations is ${}^nP_r = {}^nC_r \,.\, r!$.

Thus
$$ {}^nC_r = \frac{{}^nP_r}{r!} = \frac{n!}{r!\,(n-r)!}$$

Also
$$ {}^nC_{n-r} = {}^nC_r$$

e.g.,
$$ {}^{25}P_4 = 25 \times 24 \times 23 \times 22;\ {}^{25}C_{21} = {}^{25}C_4 = \frac{25 \times 24 \times 23 \times 22}{4 \times 3 \times 2 \times 1}.$$

Example 26.1. *In how many ways can one make a first, second, third and fourth choice among 12 firms leasing construction equipment.* (*J.N.T.U., 2003*)

Solution. First choice can be made from any of the 12 firms. Thereafter the second choice can be made from among the remaining 11 firms. Then the third choice can be made from the remaining 10 firms and the fourth choice can be made from the 9 firms.

Thus from the *principle of counting*, the number of ways in which first, second, third and fourth choice can be affected = 12 × 11 × 10 × 9 = 11880.

Example 26.2. *Find the number of permutations of all the letters of the word (i) Committee (ii) Engineering.*

Solution. (*i*) $n = 9, n_1 (m, m) = 2, n_2 (t, t) = 2, n_3 (e, e) = 2$

$$\therefore \text{ no. of permutations } = \frac{n!}{n_1!\,.\,n_2!\,.\,n_3!} = \frac{9!}{2!\,.\,2!\,.\,2!} = 45360.$$

(*ii*) $n = 11, n_1(e's) = 3, n_2 (g, g) = 2, n_3 (i, i) = 2, n_4 (n's) = 3$

$$\therefore \text{ no. of permutations } = \frac{11!}{3!\,2!\,2!\,3!} = 277200.$$

Example 26.3. *From six engineers and five architects a committee is to be formed having three engineers and two architects. How many different committees can be formed if (i) there is no restriction. (ii) two particular engineers must be included. (iii) one particular architect must be excluded.*

Solution. (*i*) Number of committees $^6C_3 \times {}^5C_2 = \frac{6\,.\,5\,.\,4}{3\,.\,2\,.\,1} \times \frac{5\,.\,4}{2\,.\,1} = 200.$

(*ii*) Here we have to choose one engineer from the remaining four engineers.

$$\therefore \text{ no. of committees} = {}^4C_1 \times {}^5C_2 = 4 \times \frac{5\,.\,4}{2\,.\,1} = 40$$

(*iii*) Here we have to choose two architects from the remaining four architects.

$$\therefore \text{ no. of committees} = {}^6C_3 \times {}^4C_2 = \frac{6\,.\,5\,.\,4}{3\,.\,2\,.\,1} \times \frac{4\,.\,3}{2\,.\,1} = 120.$$

PROBLEMS 26.1

1. If a test consists of 12 true-false questions, in how many different ways can a student make the test paper with one answer to each question. (*J.N.T.U., 2003*)
2. How many 4-digit numbers can be formed from the six digits 2, 3, 5, 6, 7 and 9, without repetition ? How many of these are less than 500 ?
3. A student has to answer 9 out of 12 questions. How many choices has he (*i*) if he must answer first two questions (*ii*) if he must answer at least four of the first five questions.
4. How many car number plates can be made if each plate contains two different letters followed by three different digits ? Solve the problem (*a*) with repetitions and (*b*) without repetitions.

26.2 (I) BASIC TERMINOLOGY

(*i*) **Exhaustive events.** A set of events is said to be *exhaustive*, if it includes all the possible events. For example, in tossing a coin there are two exhaustive cases either head or tail and there is no third possibility.

(*ii*) **Mutually exclusive events.** If the occurrence of one of the events procludes the occurrence of all other, then such a set of events is said to be *mutually exclusive*. Just as tossing a coin, either head comes up or the tail and both can't happen at the same time, *i.e.*, these are two mutually exclusive cases.

(*iii*) **Equally likely events.** If one of the events cannot be expected to happen in preference to another then such events are said to be *equally likely*. For instance, in tossing a coin, the coming of the head or the tail is equally likely.

Thus when a *die** is thrown, the turning up of the six different faces of the die are exhaustive, mutually exclusive and equally likely.

(*iv*) **Odds in favour of an event.** If the number of ways favourable to an event A is m and the number of ways not favourable to A is n then *odds in favour of* $A = m/n$ and *odds against* $A = n/m$.

(2) Definition of probability. *If there are n exhaustive, mutually exclusive and equally likely cases of which m are favourable to an event A, then probability (p) of the happening of A is*

$$\mathbf{P(A) = m/n.}$$

As there are $n - m$ cases in which A will not happen (denoted by A'), *the chance of A not happening is q or P (A′) so that*

$$q = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - p$$

i.e., $P(A') = 1 - P(A)$ so that $\mathbf{P(A) + P(A') = 1,}$

i.e., if an event is certain to happen then its probability is unity, while if it is certain not to happen, its probability is zero.

Obs. This definitions of probability fails when
(*i*) number of outcomes is infinite (not exhaustive) and (*ii*) outcomes are not equally likely.

(3) Statistical (or Empirical) definition of probability. *If in n trials, an event A happens m times, then the probability (p) of happening of A is given by*

$$p = P(A) = \underset{n \to \infty}{Lt} \frac{m}{n}$$

Example 26.4. *Find the chance of throwing (a) four, (b) an even number with an ordinary six faced die.*

Solution. (*a*) There are six possible ways in which the die can fall and of these there is only one way of throwing 4. Thus the required chance $= \frac{1}{6}$.

(*b*) There are six possible ways in which the die can fall. Of these there are only 3 ways of getting 2, 4 or 6. Thus the required chance $= 3/6 = \frac{1}{2}$.

Example 26.5. *What is the chance that a leap year selected at random will contain 53 Sundays ?* (*Madras, 2003*)

Solution. A leap year consists of 366 days, so that there are 52 full weeks (and hence 52 Sundays) and two extra days. These two days can be (*i*) Monday, Tuesday (*ii*) Tuesday, Wednesday, (*iii*) Wednesday, Thursday (*iv*) Thursday, Friday (*v*) Friday, Saturday (*vi*) Saturday, Sunday (*vii*) Sunday, Monday.

Of these 7 cases, the last two are favourable and hence the required probability $= \frac{2}{7}$.

Example 26.6. *A five figure number is formed by the digits 0, 1, 2, 3, 4 without repetition. Find the probability that the number formed is divisible by 4.*

Solution. The five digits can be arranged in 5! ways, out of which 4! will begin with zero.

∴ total number of 5-figure numbers formed $= 5! - 4! = 96$.

Those numbers formed will be divisible by 4 which will have two extreme right digits divisible by 4, *i.e.*, numbers ending in 04, 12, 20, 24, 32, 40.

Now numbers ending in $04 = 3! = 6$, numbers ending in $12 = 3! - 2! = 4$,
numbers ending in $20 = 3! = 6$, numbers ending in $24 = 3! - 2! = 4$,
numbers ending in $32 = 3! - 2! = 4$, and numbers ending in $40 = 3! = 6$.

[The numbers having 12, 24, 32 in the extreme right are $(3! - 2!)$ since the numbers having zero on the extreme left are to excluded.]

---

* *Die* is a small cube. Dots 1, 2, 3, 4, 5, 6 are marked on its six faces. The outcome of throwing a die is the number of dots on its upper face.

$\therefore$ total number of favourable ways = 6 + 4 + 6 + 4 + 4 + 6 = 30.

Hence the required probability $= \frac{30}{96} = \frac{5}{16}$.

Example 26.7. *A bag contains 40 tickets numbered 1, 2, 3, ... 40, of which four are drawn at random and arranged in ascending order ($t_1 < t_2 < t_3 < t_4$). Find the probability of t_3 being 25?*

Solution. Here exhaustive number of cases $= {}^{40}C_4$

If $t_3 = 25$, then the tickets t_1 and t_2 must come out of 24 tickets numbered 1 to 24. This can be done in ${}^{24}C_2$ ways.

Then t_4 must come out of the 15 tickets (numbering 25 to 40) which can be done in ${}^{15}C_1$ ways.

$\therefore$ favourable number of cases $= {}^{24}C_2 \times {}^{15}C_1$

Hence the probability of t_3 being 25 $= \frac{{}^{24}C_2 \times {}^{15}C_1}{{}^{40}C_4} = \frac{414}{9139}$.

Example 26.8. *An urn contains 5 red and 10 black balls. Eight of them are placed in another urn. What is the chance that the latter then contains 2 red and 6 black balls ?*

Solution. The number of ways in which 8 balls can be drawn out of 15 is ${}^{15}C_8$.

The number of ways of drawing 2 red balls is ${}^5C_2$ and corresponding to each of these ${}^5C_2$ ways of drawing a red ball, there are ${}^{10}C_6$ ways of drawing 6 black balls.

$\therefore$ the total number of ways in which 2 red and 6 black balls can be drawn is ${}^5C_2 \times {}^{10}C_6$.

$\therefore$ the required probability $= \frac{{}^5C_2 \times {}^{10}C_6}{{}^{15}C_8} = \frac{140}{429}$.

Example 26.9. *A committee consists of 9 students two of which are from 1st year, three from 2nd year and four from 3rd year. Three students are to be removed at random. What is the chance that (i) the three students belong to different classes, (ii) two belong to the same class and third to the different class, (iii) the three belong to the same class ?* (V.T.U., 2002 S)

Solution. (*i*) The total number of ways of choosing 3 students out of 9 is ${}^9C_3$, *i.e.*, 84.

A student can be removed from 1st year students in 2 ways, from 2nd year in 3 ways and from 3rd year in 4 ways, so that the total number of ways of removing three students, one from each group is 2 × 3 × 4.

Hence the required chance $= \frac{2 \times 3 \times 4}{{}^9C_3} = \frac{24}{84} = \frac{2}{7}$.

(*ii*) The number of ways of removing two from 1st year students and one from others

$$= {}^2C_2 \times {}^7C_1.$$

The number of ways of removing two from 2nd year students and one from others

$$= {}^3C_2 \times {}^6C_1.$$

The number of ways of removing 2 from 3rd year students and one from others

$$= {}^4C_2 \times {}^5C_1.$$

$\therefore$ the total number of ways in which two students of the same class and third from the others may be removed

$$= {}^2C_2 \times {}^7C_1 + {}^3C_2 \times {}^6C_1 + {}^4C_2 \times {}^5C_1 = 7 + 18 + 30 = 55.$$

Hence, the required chance $= \frac{55}{84}$.

(*iii*) Three students can be removed from 2nd year group in ${}^3C_3$, *i.e.* 1 way and from 3rd year group in ${}^4C_3$, *i.e.*, 4 ways.

$\therefore$ the total number of ways in which three students belong to the same class = 1 + 4 = 5.

Hence the required chance $= \frac{5}{84}$.

Example 26.10. *A has one share in a lottery in which there is 1 prize and 2 blanks ; B has three shares in a lottery in which there are 3 prizes and 6 blanks ; compare the probability of A's success to that of B's success.*

Solution. A can draw a ticket in $^3C_1 = 3$ ways.

The number of cases in which A can get a prize is clearly 1.

$\therefore$ the probability of A's success $= \dfrac{1}{3}$.

Again B can draw a ticket in $^9C_3 = \dfrac{9 \cdot 8 \cdot 7}{3 \cdot 2 \cdot 1} = 84$ ways.

The number of ways in which B gets all blanks $= {}^6C_3 = \dfrac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 20$

$\therefore$ the number of ways of getting a prize $= 84 - 20 = 64$.

Thus the probability of B's success $= 64/84 = 16/21$.

Hence A's probability of success : B's probability of success $= \dfrac{1}{3} : \dfrac{16}{21} = 7 : 16$.

26.3 PROBABILITY AND SET NOTATIONS

(1) Random experiment. Experiments which are performed essentially under the same conditions and whose results cannot be predicted are known as *random experiments. e.g.,* Tossing a coin or rolling a die are random experiments.

(2) Sample space. The set of all possible outcomes of a random experiment is called *sample space* for that experiment and is denoted by S.

The elements of the sample space S are called the *sample points.*

e.g., On tossing a coin, the possible outcomes are the head (H) and the tail (T). Thus $S = \{H, T\}$.

(3) Event. The outcome of a random experiment is called an *event*. Thus every subset of a sample space S is an *event*.

The null set ϕ is also an event and is called an *impossible event. Probability of an impossible event* is zero *i.e.*, $P(\phi) = 0$.

(4) Axioms

(*i*) *The numerical value of probability lies between 0 and 1.*

i.e., for any event A of S, $0 \le P(A) \le 1$.

(*ii*) *The sum of probabilities of all sample events is unity i.e.,* $P(S) = 1$.

(*iii*) *Probability of an event made of two or more sample events is the sum of their probabilities.*

(5) Notations

(*i*) *Probability of happening of events* **A** *or* **B** is written as $P(A + B)$ or $P(A \cup B)$.

(*ii*) *Probability of happening of* **both the events A and B** is written as $P(AB)$ or $P(A \cap B)$.

(*iii*) '*Event A implies* ($\Rightarrow$) *event B*' is expressed as $A \subset B$.

(*iv*) '*Events A and B are mutually exclusive*' is expressed as $A \cap B = \phi$.

(6) *For any two events A and B,*

$$\boldsymbol{P(A \cap B') = P(A) - P(A \cap B)}$$

Proof. From Fig. 26.1,

$$(A \cap B') \cup (A \cap B) = A$$

$$\therefore \quad P[(A \cap B') \cup (A \cap B)] = P(A)$$

or $$P(A \cap B') + P(A \cap B) = P(A)$$

or $$P(A \cap B') = P(A) - P(A \cap B)$$

Similarly, $P(A' \cap B) = P(B) - P(A \cap B)$

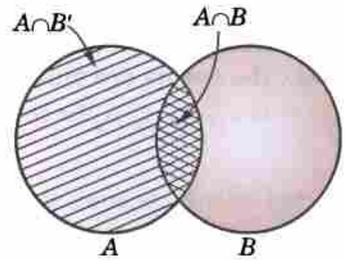

Fig. 26.1

26.4 ADDITION LAW OF PROBABILITY *or* THEOREM OF TOTAL PROBABILITY

(1) *If the probability of an event A happening as a result of a trial is P(A) and the probability of a* **mutually exclusive** *event B happening is P(B), then the probability of* **either** *of the events happening as a result of the trial is* **$P(A + B)$ or $P(A \cup B) = P(A) + P(B)$.**

Proof. Let n be the total number of equally likely cases and let m_1 be favourable to the event A and m_2 be favourable to the event B. Then the number of cases favourable to A or B is $m_1 + m_2$. Hence the probability of A or B happening as a result of the trial

$$= \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

(2) *If A, B, are any two events* **(not mutually exclusive),** *then*

$$\boldsymbol{P(A + B) = P(A) + P(B) - P(AB)}$$

or

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

If the events A and B are any two events then, there are some outcomes which favour both A and B. If m_3 be their number, then these are included in both m_1 and m_2. Hence the total number of outcomes favouring either A or B or both is

$$m_1 + m_2 - m_3.$$

Thus the probability of occurrence of A or B or both

$$= \frac{m_1 + m_2 - m_3}{n} = \frac{m_1}{n} + \frac{m_2}{n} - \frac{m_3}{n}$$

Hence $\quad P(A + B) = P(A) + P(B) - P(AB)$

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

Obs. *When A and B are mutually exclusive* $P(AB)$ or $P(A \cap B) = 0$ and we get

$$P(A + B) \text{ or } P(A \cup B) = P(A) + P(B).$$

In general, for a number of **mutually exclusive** events $A_1, A_2, \ldots A_n$, we have

$$P(A_1 + A_2 + \ldots + A_n) \text{ or } P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

(3) *If A, B, C are any three events, then*

$$\boldsymbol{P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC).}$$

or

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

Proof. Using the above result for any two events, we have

$$P(A \cup B \cup C) = P[(A \cup B) \cup C]$$

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

$$= [P(A) + P(B) - P(A \cap B)] + P(C) - P[(A \cap C) \cup (B \cap C)] \qquad \text{(Distributive Law)}$$

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

$$[\because \ (A \cap C) \cap (B \cap C) = A \cap B \cap C]$$

$$= P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C) \ [\because \ A \cap C = C \cap A.]$$

Example 26.11. *In a race, the odds in favour of the four horses H_1, H_2, H_3, H_4 are 1 : 4, 1 : 5, 1 : 6, 1 : 7 respectively. Assuming that a dead heat is not possible, find the chance that one of them wins the race.*

Solution. Since it is not possible for all the horses to cover the same distance in the same time (a dead heat), the events are mutually exclusive.

If p_1, p_2, p_3, p_4 be the probabilities of winning of the horses H_1, H_2, H_3, H_4 respectively, then

$$p_1 = \frac{1}{1+4} = \frac{1}{5} \qquad [\because \text{ Odds in favour of } H_1 \text{ are } 1:4]$$

and

$$p_2 = \frac{1}{6}, p_3 = \frac{1}{7}, p_4 = \frac{1}{8}.$$

Hence the chance that one of them wins $= p_1 + p_2 + p_3 + p_4$

$$= \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} = \frac{533}{840}.$$

Example 26.12. *A bag contains 8 white and 6 red balls. Find the probability of drawing two balls of the same colour.*

Solution. Two balls out of 14 can be drawn in ${}^{14}C_2$ ways which is the total number of outcomes.

Two white balls out of 8 can be drawn in ${}^{8}C_2$ ways. Thus the probability of drawing 2 white balls

$$= \frac{{}^{8}C_2}{{}^{14}C_2} = \frac{28}{91}$$

Similarly 2 red balls out of 6 can be drawn in ${}^{6}C_2$ ways. Thus the probability of drawing 2 red balls

$$= \frac{{}^{6}C_2}{{}^{14}C_2} = \frac{15}{91}.$$

Hence the probability of drawing 2 balls of the same colour (either both white or both red)

$$= \frac{28}{91} + \frac{15}{91} = \frac{43}{91}.$$

Example 26.13. *Find the probability of drawing an ace or a spade or both from a deck of cards* ?*

Solution. The probability of drawing an ace from a deck of 52 cards = 4/52.

Similarly the probability of drawing a card of spades = 13/52, and the probability of drawing an ace of spades = 1/52.

Since the two events (*i.e.*, a card being an ace and a card being of spades) are not mutually exclusive, therefore, the probability of drawing an ace or a spade

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{4}{13}.$$

26.5 (1) INDEPENDENT EVENTS

Two events are said to be *independent*, if happening or failure of one does not affect the happening or failure of the other. Otherwise the events are said to be *dependent*.

For two dependent events A and B, the symbol $P(B/A)$ denotes the probability of occurrence of B, when A has already occurred. It is known as the **conditional probability** and is read as a '*probability of B given A*'.

(2) Multiplication law of probability *or* **Theorem of compound probability.** *If the probability of an event A happening as a result of trial is P(A) and after A has happened the probability of an event B happening as a result of another trial* (*i.e.*, **conditional probability of B given A**) *is $P(B/A)$, then the probability of* **both** *the events A and B happening as a result of two trials is* **$P(AB)$ or $P(A \cap B) = P(A) . P(B/A)$.**

Proof. Let n be the total number of outcomes in the first trial and m be favourable to the event A so that $P(A) = m/n$.

Let n_1 be the total number of outcomes in the second trial of which m_1 are favourable to the event B so that $P(B/A) = m_1/n_1$.

Now each of the n outcomes can be associated with each of the n_1 outcomes. So the total number of outcomes in the combined trial is nn_1. Of these mm_1 are favourable to both the events A and B. Hence

$$P(AB) \text{ or } P(A \cap B) = \frac{mm_1}{nn_1} = P(A) . P(B/A).$$

Similarly, the *conditional probability of A given B is $P(A/B)$*.

$\therefore \quad P(AB)$ or $P(A \cap B) = P(B) . P(A/B)$

Thus $\quad P(A \cap B) = P(A) . P(B/A) = P(B) . P(A/B)$.

(3) *If the events A and B are* **independent,** *i.e.*, if the happening of B does not depend on whether A has happened or not, then $P(B/A) = P(B)$ and $P(A/B) = P(A)$.

$\therefore \quad$ **$P(AB)$ or $P(A \cap B) = P(A) . P(B)$.**

In general, $\quad P(A_1A_2...A_n)$ or $P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1) . P(A_2)... . P(A_n)$.

* *Cards* : A pack of cards consists of four suits *i.e.*, Hearts, Diamonds, Spades and Clubs. Each suit has 13 cards : an Ace, a King, a Queen, a Jack and nine cards numbered 2, 3, 4, ..., 10. Hearts and Diamonds are *red* while Spades and Clubs are *black*.

Cor. If p_1, p_2 be the probabilities of happening of two independent events, then

(*i*) the probability that the first event happens and the second fails is $p_1(1-p_2)$.

(*ii*) the probability that both events fail to happen is $(1-p_1)(1-p_2)$.

(*iii*) the probability that at least one of the events happens is

$1-(1-p_1)(1-p_2)$. This is commonly known as their **cumulative probability.**

In general, if $p_1, p_2, p_3, \ldots p_n$ be the chances of happening of n independent events, then their cumulative probability (i.e., the chance that at least one of the events will happen) is

$$1-(1-p_1)(1-p_2)(1-p_3)\ldots(1-p_n).$$

Example 26.14. *Two cards are drawn in succession from a pack of 52 cards. Find the chance that the first is a king and the second a queen if the first card is (i) replaced, (ii) not replaced.*

Solution. (*i*) The probability of drawing a king $= \frac{4}{52} = \frac{1}{13}$.

If the card is replaced, the pack will again have 52 cards so that the probability of drawing a queen is 1/13.

The two events being independent, the probability of drawing both cards in succession $= \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}$.

(*ii*) The probability of drawing a king $= \frac{1}{13}$.

If the card is not replaced, the pack will have 51 cards only so that the chance of drawing a queen is 4/51.

Hence the probability of drawing both cards $= \frac{1}{13} \times \frac{4}{51} = \frac{4}{663}$.

Example 26.15. *A pair of dice is tossed twice. Find the probability of scoring 7 points (a) once, (b) at least once (c) twice.* *(Kurukshetra, 2009 S ; V.T.U., 2004)*

Solution. In a single toss of two dice, the sum 7 can be obtained as (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1) *i.e.*, in 6 ways, so that the probability of getting 7 = 6/36 = 1/6.

Also the probability of not getting 7 = 1 – 1/6 = 5/6.

(*a*) The probability of getting 7 in the first toss and not getting 7 in the second toss = 1/6 × 5/6 = 5/36.

Similarly, the probability of not getting 7 in the first toss and getting 7 in the second toss = 5/6 × 1/6 = 5/36.

Since these are mutually exclusive events, addition law of probability applies.

$\therefore$ required probability $= \frac{5}{36} + \frac{5}{36} = \frac{5}{18}$.

(*b*) The probability of not getting 7 in either toss $= \frac{5}{6} \times \frac{5}{6} = \frac{25}{36}$

$\therefore$ the probability of getting 7 at least once $= 1 - \frac{25}{36} = \frac{11}{36}$.

(*c*) The probability of getting 7 twice $= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

Example 26.16. *There are two groups of objects : 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 the 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*

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

Prob. of selecting an engg. subject from first group $= \frac{3}{8}$

$\therefore$ Prob of selecting an engg. subject from first group on turning up 3 or 5

$$= \frac{1}{3} \times \frac{3}{8} = \frac{1}{8} \qquad \ldots(i)$$

Now prob. of not turning 3 or 5 $= 1 - \frac{1}{3} = \frac{2}{3}$.

Prob. of selecting an engg. subject from second group $= \frac{5}{8}$

$\therefore$ prob. of selecting an engg. subject from second group on turning up 3 or 5

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

Thus the prob. of selecting an engg. subject

$$= \frac{1}{8} + \frac{5}{12} = \frac{13}{24}. \quad \text{[From (i) and (ii)]}$$

Example 26.17. *A box A contains 2 white and 4 black balls. Another box B contains 5 white and 7 black balls. A ball is transferred from the box A to the box B. Then a ball is drawn from the box B. Find the probability that it is white.* (V.T.U., 2004)

Solution. The probability of drawing a white ball from box B will depend on whether the transferred ball is black or white.

If a black ball is transferred, its probability is 4/6. There are now 5 white and 8 black balls in the box B. Then the probability of drawing white ball from box B is $\frac{5}{13}$.

Thus the probability of drawing a white ball from urn B, if the transferred ball is black

$$= \frac{4}{6} \times \frac{5}{13} = \frac{10}{39}.$$

Similarly the probability of drawing a white ball from urn B, if the transferred ball is white

$$= \frac{2}{6} \times \frac{6}{13} = \frac{2}{13}.$$

Hence required probability $= \frac{10}{39} + \frac{2}{13} = \frac{16}{39}$.

Example 26.18. *(a) A biased coin is tossed till a head appears for the first time. What is the probability that the number of required tosses is odd.* (Mumbai, 2006)

(b) Two persons A and B toss an unbiased coin alternately on the understanding that the first who gets the head wins. If A starts the game, find their respective chances of winning. (Madras, 2000 S)

Solution. (*a*) Let p be the probability of getting a head and q the probability of getting a tail in a single toss, so that $p + q = 1$.

Then probability of getting head on an odd toss

= Probability of getting head in the 1st toss
+ Probability of getting head in the 3rd toss
+ Probability of getting head in the 5th toss + ... ∞

$$= p + qqp + qqqqp + \ldots \infty$$

$$= p(1 + q^2 + q^4 + \ldots) = p \cdot \frac{1}{1 - q^2} \quad (q < 1)$$

$$= p \cdot \frac{1}{(1-q)(1+q)} = p \cdot \frac{1}{p(1+q)} = \frac{1}{1+q}.$$

(*b*) Probability of getting a head = 1/2. Then A can win in 1st, 3rd, 5th, ... throws.

$$\therefore \text{ the chances of } A\text{'s winning } = \frac{1}{2} + \left(\frac{1}{2}\right)^2 \frac{1}{2} + \left(\frac{1}{2}\right)^4 \frac{1}{2} + \left(\frac{1}{2}\right)^6 \frac{1}{2} + \cdots$$

$$= \frac{1/2}{1 - (1/2)^2} = \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}.$$

Hence the chance of B's winning $= 1 - 2/3 = 1/3$.

Example 26.19. *Two cards are selected at random from 10 cards numbered 1 to 10. Find the probability p that the sum is odd, if*

(i) the two cards are drawn together.

(ii) the two cards are drawn one after the other without replacement.

(iii) the two cards are drawn one after the other with replacement. (J.N.T.U., 2003)

Solution. (*i*) Two cards out of 10 can be selected in ${}^{10}C_2 = 45$ ways. The sum is odd if one number is odd and the other number is even. There being 5 odd numbers (1, 3, 5, , 9) and 5 even numbers (2, 4, 6, 8, 10), an odd and an even number is chosen in 5 × 5 = 25 ways.

Thus $$p = \frac{25}{45} = \frac{5}{9}.$$

(*ii*) Two cards out of 10 can be selected one after the other *without replacement* in 10 × 9 = 90 ways.

An odd number is selected in 5 × 5 = 25 ways and an even number in 5 × 5 = 25 ways

Thus $$p = \frac{25+25}{90} = \frac{5}{9}.$$

(*iii*) Two cards can be selected one after the other *with replacement* in 10 × 10 = 100 ways.

An odd number in selected in 5 × 5 = 25 ways and an even number in 5 × 5 = 25 ways.

Thus $$p = \frac{25+25}{100} = \frac{1}{2}.$$

Example 26.20. *Given $P(A) = 1/4$, $P(B) = 1/3$ and $P(A \cup B) = 1/2$, evaluate $P(A/B)$, $P(B/A)$, $P(A \cap B')$ and $P(A/B')$.*

Solution. (*i*) Since $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\therefore \quad \frac{1}{2} = \frac{1}{4} + \frac{1}{3} - P(A \cap B) \text{ or } P(A \cap B) = \frac{1}{12}$$

Thus $$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/12}{1/3} = \frac{1}{4}.$$

(*ii*) $$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/12}{1/4} = \frac{1}{3}.$$

(*iii*) $$P(A \cap B') = P(A) - P(A \cap B) = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}.$$

(*iv*) $$P(A/B') = \frac{P(A \cap B')}{P(B')} = \frac{1/6}{1 - P(B)} = \frac{1/6}{1 - 1/3} = \frac{1}{4}.$$

Example 26.21. *The odds that a book will be reviewed favourably by three independent critics are 5 to 2, 4 to 3 and 3 to 4. What is the probability that of the three reviews, a majority will be favourable.* (V.T.U., 2003 S)

Solution. The probability that the book shall be reviewed favourably by first critic is 5/7, by second 4/7 and by third 3/7.

A majority of the three reviews will be favourable when two or three are favourable.

∴ prob. that the first two are favourable and the third unfavourable

$$= \frac{5}{7} \times \frac{4}{7} \times \left(1 - \frac{3}{7}\right) = \frac{80}{343}$$

Prob. that the first and third are favourable and second unfavourable

$$= \frac{5}{7} \times \frac{3}{7} \times \left(1 - \frac{4}{7}\right) = \frac{45}{343}$$

Prob. that the second and third are favourable and the first unfavourable

$$= \frac{4}{7} \times \frac{3}{7} \times \left(1 - \frac{5}{7}\right) = \frac{24}{343}$$

Finally, prob. that all the three are favourable $= \frac{5}{7} \times \frac{4}{7} \times \frac{3}{7} = \frac{60}{343}$

Since they are mutually exclusive events, the required prob.

$$= \frac{80}{343} + \frac{45}{343} + \frac{24}{343} + \frac{60}{343} = \frac{209}{343}.$$

Example 26.22. *I can hit a target 3 times in 5 shots, B 2 times in 5 shots and C 3 times in 4 shots. They fire a volley. What is the probability that (i) two shots hit, (ii) atleast two shots hit ?*

(A.M.I.E.T.E., 2003 ; Madras, 2000 S)

Solution. Prob. of A hitting the target = 3/5, prob. of B hitting the target = 2/5

Prob. of C hitting the target = 3/4.

(*i*) In order that two shots may hit the target, the following cases must be considered :

$$p_1 = \text{Chance that } A \text{ and } B \text{ hit and } C \text{ fails to hit} = \frac{3}{5} \times \frac{2}{3} \times \left(1 - \frac{3}{4}\right) = \frac{6}{100}$$

$$p_2 = \text{Chance that } B \text{ and } C \text{ hit and } A \text{ fails to hit} = \frac{2}{5} \times \frac{3}{4} \times \left(1 - \frac{3}{5}\right) = \frac{12}{100}$$

$$p_3 = \text{Chance that } C \text{ and } A \text{ hit and } B \text{ fails to hit} = \frac{3}{4} \times \frac{3}{5} \times \left(1 - \frac{2}{5}\right) = \frac{27}{100}$$

Since these are mutually exclusive events, the probability that any 2 shots hit

$$= p_1 + p_2 + p_3 = \frac{6}{100} + \frac{12}{100} + \frac{27}{100} = 0.45.$$

(*ii*) In order that at least two shots may hit the target, we must also consider the case of all A, B, C hitting the target [in addition to the three cases of (*i*)] for which

$$p_4 = \text{chance that } A, B, C \text{ all hit} = \frac{3}{5} \times \frac{2}{5} \times \frac{3}{4} = \frac{18}{100}$$

Since all these are mutually exclusive events, the probability of *atleast* two shots hit

$$= p_1 + p_2 + p_3 + p_4 = \frac{6}{100} + \frac{12}{100} + \frac{27}{100} + \frac{18}{100} = 0.63.$$

Example 26.23. *A problem in mechanics is given to three students A, B, and C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{4}$ respectively. What is the probability that the problem will be solved.* *(V.T.U., 2004)*

Solution. The probability that A can solve the problem is 1/2.

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

Similarly the probabilities that B and C cannot solve the problem are $1 - \frac{1}{3}$ and $1 - \frac{1}{4}$.

$\therefore$ the probability that A, B and C cannot solve the problem is $\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{4}\right)$.

Hence the probability that the problem will be solved, *i.e.*, at least one student will solve it

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

Example 26.24. *The students in a class are selected at random, one after the other, for an examination. Find the probability p that the boys and girls in the class alternate if*

(*i*) *the class consists of 4 boys and 3 girls.*

(*ii*) *the class consists of 3 boys and 3 girls.* *(J.N.T.U., 2003)*

Solution. (*i*) As there are 7 students in the class, the first examined must be a boy.

$\therefore$ prob. that first is a boy $= \frac{4}{7}$

Then the prob. that the second is a girl $= \frac{3}{6}$.

$\therefore$ prob. of the next boy $= \frac{3}{5}$

Similarly the prob. that the fourth is a girl $= \frac{2}{4}$,

the prob. that the fifth is a boy $= \frac{2}{3}$,

the prob. that the sixth is a girl $= \frac{1}{2}$

and the last is a boy $= \frac{1}{1}$.

Thus $$p = \frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{35}.$$

(*ii*) The first student is a boy and the first student is a girl are two mutually exclusive cases. If the first student is a boy, then the probability p_1 that the students alternate is

$$p_1 = \frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{20}.$$

If the first student is a girl, then the probability p_2 that the students alternate is

$$p_2 = \frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{20}.$$

Thus the required prob. $p = p_1 + p_2 = \frac{1}{20} + \frac{1}{20} = \frac{1}{10}$.

Example 26.25. (*Huyghen's problem*) *A and B throw alternately with 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. If A begins, find his chance of winning.*

(*Madras, 2006 ; J.N.T.U., 2003*)

Solution. The sum 6 can be obtained as follows : (1, 5), (2, 4), (3, 3), (4, 2), (5, 1), *i.e.*, in 5 ways.

The probability of A's throwing 6 with 2 dice is $\frac{5}{36}$.

$\therefore$ the probability of A's not throwing 6 is 31/36.

Similarly the probability of B's throwing 7 is 6/36, *i.e.*, $\frac{1}{6}$.

$\therefore$ the probability of B's not throwing 7 is 5/6.

Now A can win if he throws 6 in the first, third, fifth, seventh etc. throws.

$\therefore$ the chance of A's wining

$$= \frac{5}{36} + \frac{31}{36}\times\frac{5}{6}\times\frac{5}{36} + \frac{31}{36}\times\frac{5}{6}\times\frac{31}{36}\times\frac{5}{6}\times\frac{5}{36} + \cdots$$

$$= \frac{5}{36}\left[1 + \left(\frac{31}{36}\times\frac{5}{6}\right) + \left(\frac{31}{36}\times\frac{5}{6}\right)^2 + \left(\frac{31}{36}\times\frac{5}{6}\right)^3 + \cdots\right]$$

$$= \frac{5}{36}\cdot\frac{1}{1-(31/36)\times(5/6)} = \frac{5}{36}\times\frac{36\times 6}{61} = \frac{30}{61}.$$

PROBLEMS 26.2

1. (*i*) Given $P(A) = 1/2$, $P(B) = 1/3$ and $P(AB) = 1/4$, find the value $P(A + B)$. (*Burdwan, 2003*)

(*ii*) Let A and B be two events with $P(A) = 1/2$, $P(B) = 1/3$ and $P(A \cap B) = 1/4$. Find $P(A/B)$, $P(A \cup B)$, $P(A'/B')$.

(*Kurukshetra, 2009 ; V.T.U., 2003 S*)

2. In a single throw with two dice, what is the chance of throwing
(*a*) two aces ? (*b*) 7 ? Is this probability the same as that for getting 7 in two throws of a single die ?

3. Compare the chances of throwing 4 with one dice, 8 with two dice and 12 with three dice.

4. Find the probability that a non-leap year should have 53 Saturdays ? (*Madras, 2003*)

5. When a coin is tossed four times, find the probability of getting (*i*) exactly one head, (*ii*) at most three heads and (*iii*) at least two heads ? (*V.T.U., 2000 S*)

6. Ten coins are thrown simultaneously. Find the probability of getting at least seven heads. (*P.T.U., 2003*)

7. If all the letters of word 'ENGINEER' be written at random, what is the probability that all the latters E are found together.

8. A ten digit number is formed using the digits from zero to nine, every digit being used only once. Find the probability that the number is divisible by 4.

9. Four cards are drawn from a pack of 52 cards. What is the chance that
(*i*) no two cards are of equal value ? (*ii*) each belongs to a different suit ?

10. Suppose 5 cards are drawn at random from a pack of 52 cards. If all cards are red what is the probability that all of them are hearts? (*Mumbai, 2005*)

11. Out of 50 rare books, 3 of which are especially valuable, 5 are stolen at random by a thief. What is the probability that
(*a*) none of the 3 is included ? (*b*) 2 of the 3 are included ?

12. Five men in a company of twenty are graduates. If 3 men are picked out of 20 at random, what is the probability that
(*a*) they are all graduates ? (*b*) at least one is graduate ?

13. From 20 tickets marked from 1 to 20, one ticket is drawn at random. Find the probability that it is marked with a multiple of 3 or 5.

14. Five balls are drawn from a bag containing 6 white and 4 black balls. What is the chance that 3 white and 2 black balls are drawn ?

15. The probability of n independent events are $p_1, p_2, p_3, \ldots p_n$. Find the probability that at least one of the events will happen. Use this result to find the chance of getting at least one six in a throw of 4 dice.

16. Find the probability of drawing 4 white balls and 2 black balls without replacement from a bag containing 1 red, 4 black and 6 white balls.

17. A bag contains 10 white and 15 black balls. Two balls are drawn in succession. What is the probability that one of them is black and the other white ?

18. A purse contains 2 silver and 4 copper coins and a second purse contains 4 silver and 4 copper coins. If a coin is selected at random from one of the two purses, what is the probability that it is a silver coin ? (*Osmania 2002*)

19. A box I contains 5 white balls and 6 black balls. Another box II contains 6 white balls and 4 black balls. A box is selected at random and then a ball is drawn from it : (*i*) what is the probability that the ball drawn will be white ? (*ii*) Given that the ball drawn is white, what is the probability that it came from box I. (*Mumbai, 2006*)

20. A party of n persons take their seats at random at a round table ; find the probability that two specified persons do not sit together.

21. A speaks the truth in 75% cases, and B in 80% of the cases. In what percentage of cases, are they likely to contradict each other in stating the same fact ? (*V.T.U., 2002 S*)

22. The probability that Sushil will solve a problem is 1/4 and the probability that Ram will solve it is 2/3. If Sushil and Ram work independently, what is the probability that the problem will be solved by (*a*) both of them, (*b*) at least one of them ?

23. A student takes his examination in four subjects, P, Q, R, S. He estimates his chances of passing in P as 4/5, in Q as 3/4, in R as 5/6 and in S as 2/3. To qualify, he must pass in P and at least two other subjects. What is the probability that he qualifies ? (*Madras, 2000 S*)

24. The probability that a 50 year old man will be alive at 60 is 0.83 and the probability that a 45 year old women will be alive at 55 is 0.87. What is the probability that a man who is 50 and his wife who is 45 will both be alive 10 years hence ?

25. If on an average one birth in 80 is a case of twins, what is the probability that there will be at least one case of twins in a maternity hospital on a day when 20 births occur ?

26. Two persons A and B fire at a target independently and have a probability 0.6 and 0.7 respectively of hitting the target. Find the probability that the target is destroyed.

27. A and B throw alternately with a pair of dice. The one who throws 9 first wins. Show that the chances of their winning are 9 : 8.

26.6 BAYE'S THEOREM

An event A corresponds to a number of exhaustive events $B_1, B_2, ..., B_n$. If $P(B_i)$ and $P(A/B_i)$ are given, then

$$P(B_i/A) = \frac{P(B_i)\,P(A/B_i)}{\Sigma P(B_i)P(A/B_i)}.$$

Proof. By the multiplication law of probability,

$$P(AB_i) = P(A)\,P(B_i/A) = P(B_i)\,P(A/B_i) \qquad ...(1)$$

$$\therefore \quad P(B_i/A) = \frac{P(B_i)\,P(A/B_i)}{P(A)} \qquad ...(2)$$

Since the event A corresponds to $B_1, B_2, ..., B_n$, we have by the addition law of probability,

$$P(A) = P(AB_1) + P(AB_2) + ... + P(AB_n) = \Sigma P(AB_i) = \Sigma P(B_i)\,P(A/B_i) \qquad \text{[By (1)]}$$

Hence from (2), we have $P(B_i/A) = \dfrac{P(B_i)\,P(A/B_i)}{\Sigma P(B_i)\,P(A/B_i)}$

which is known as the *theorem of inverse probability.*

Obs. The probabilities $P(B_i)$, $i = 1, 2, ..., n$ are called *apriori probabilities* because these exist before we get any information from the experiment.

The probabilities $P(A/B_i)$, $i = 1, 2, ..., n$ are called *posteriori probabilities*, because these are found after the experiment results are known.

Example 26.26. *Three machines M_1, M_2 and M_3 produce identical items. Of their respective output 5%, 4% and 3% of items are faulty. On a certain day, M_1 has produced 25% of the total output, M_2 has produced 30% and M_3 the remainder. An item selected at random is found to be faulty. What are the chances that it was produced by the machine with the highest output?*

Solution. Let the event of drawing a faulty item from any of the machines be A, and the event that an item drawn at random was produced by M_i be B_i. We have to find $P(B_i/A)$ for which we proceed as follows :

	M_1	M_2	M_3	*Remarks*
$P(B_i)$	0.25	0.30	0.45	$\because$ sum = 1
$P(A/B_i)$	0.05	0.04	0.03	
$P(B_i)\,P(A/B_i)$	0.0125	0.012	0.0135	sum = 0.38
$P(B_i/A)$	$\frac{0.0125}{0.038}$	$\frac{0.012}{0.038}$	$\frac{0.0135}{0.038}$	by Baye's theorem

The highest output being from M_3, the required probability = 0.0135/0.038 = 0.355.

Example 26.27. *There are three bags : first containing 1 white, 2 red, 3 green balls ; second 2 white, 3 red, 1 green balls and third 3 white, 1 red, 2 green balls. Two balls are drawn from a bag chosen at random. These are found to be one white and one red. Find the probability that the balls so drawn came from the second bag.* (J.N.T.U., 2003)

Solution. Let B_1, B_2, B_3 pertain to the first, second, third bags chosen and A : the two balls are white and red.

Now $\quad P(B_1) = P(B_2) = P(B_3) = \dfrac{1}{3}$

$P(A/B_1) = P$ (a white and a red ball are drawn from first bag)

$$= ({}^1C_1 \times {}^2C_1)/{}^6C_2 = \frac{2}{15}$$

Similarly $\quad P(A/B_2) = ({}^2C_1 \times {}^3C_1)/{}^6C_2 = \dfrac{2}{5}, \; P(A/B_3) = ({}^3C_1 \times {}^1C_1)/{}^6C_2 = \dfrac{1}{5}$

By Baye's theorem,

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

$$= \frac{\frac{1}{3}\times\frac{2}{5}}{\frac{1}{3}\times\frac{2}{15}+\frac{1}{3}\times\frac{2}{5}+\frac{1}{3}\times\frac{1}{5}} = \frac{6}{11}.$$

PROBLEMS 26.3

1. In a certain college, 4% of the boys and 1% of girls are taller than 1.8 m. Further more 60% of the students are girls. If a student is selected at random and is found to be taller than 1.8 m., what is the probability that the student is a girl ?
2. In a bolt factory, machines *A*, *B* and *C* manufacture 25%, 35% and 40% of the total. Of their output 5%, 4% and 2% are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines *A*, *B* or *C* ? (*V.T.U., 2006 ; Rohtak, 2005 ; Madras, 2000 S*)
3. In a bolt factory, there are four machines *A*, *B*, *C*, *D* manufacturing 20%, 15%, 25% and 40% of the total output respectively. Of their outputs 5%, 4%, 3% and 2% in the same order are defective bolts. A bolt is chosen at random from the factory's production and is found defective. What is the probability that the bolt was manufactured by machine *A* or machine *D* ? (*Hissar, 2007 ; J.N.T.U., 2003*)
4. The contents of three urns are : 1 white, 2 red, 3 green balls ; 2 white, 1 red, 1 green balls and 4 white, 5 red, 3 green balls. Two balls are drawn from an urn chosen at random. These are found to be one white and one green. Find the probability that the balls so drawn came from the third urn. (*Kurukshetra, 2007*)

26.7 RANDOM VARIABLE

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

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

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

26.8 (1) DISCRETE PROBABILITY DISTRIBUTION

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

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

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

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

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

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

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

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

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

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

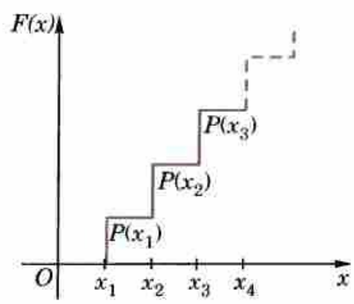

Fig. 26.2

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

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

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

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

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

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

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

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

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

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

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

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

(i) Find $P(X < 4)$, $P(X \geq 5)$, $P(3 < X \leq 6)$. (V.T.U., 2010)

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

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

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

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

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

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

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

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

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

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

(i) Find the value of the k *(ii) Evaluate $P(X < 6)$, $P(X \geq 6)$*

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

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

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

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

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

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

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

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

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

26.9 (1) CONTINUOUS PROBABILITY DISTRIBUTION

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

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

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

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

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

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

(2) Distribution function

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

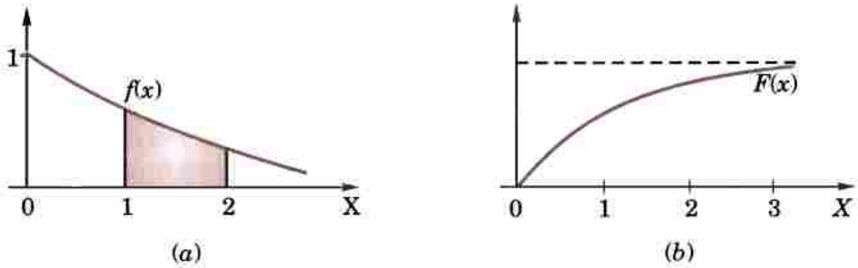

Fig. 26.3

26.10 (1) EXPECTATION

The mean value (μ) *of the probability distribution of a variate X is commonly known as its* **expectation** *and is denoted by E(X).* If $f(x)$ is the probability density function of the variate X, then

$$\sum_i x_i f(x_i) \quad \text{(discrete distribution)}$$

or $$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx \quad \text{(continuous distribution)}$$

In general, expectation of any function $\phi(x)$ is given by

$$E[\phi(x)] = \sum_i \phi(x_i) f(x_i) \quad \text{(discrete distribution)}$$

or $$E[\phi(x)] = \int_{-\infty}^{\infty} \phi(x) f(x)\,dx \quad \text{(continuous distribution)}$$

(2) **Variance** *of a distribution is given by*

$$\sigma^2 = \sum_i (x_i - \mu)^2 f(x_i) \quad \text{(discrete distribution)}$$

or $$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,dx \quad \text{(continuous distribution)}$$

where σ is the *standard deviation* of the distribution.

(3) *The* **rth moment** *about the mean* (denoted by μ_r) *is defined by*

$$\mu_r = \Sigma (x_i - \mu)^r f(x_i) \quad \text{(discrete distribution)}$$

or $$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x)\,dx \quad \text{(continuous distribution)}$$

(4) **Mean deviation from the mean** *is given by*

$$\Sigma\, | x_i - \mu |\, f(x_i) \quad \text{(discrete distribution)}$$

or by $$\int_{-\infty}^{\infty} | x - \mu |\, f(x)\,dx \quad \text{(continuous distribution)}$$

Example 26.32. *In a lottery, m tickets are drawn at a time out of n tickets numbered from 1 to n. Find the expected value of the sum of the numbers on the tickets drawn.*

Solution. Let $x_1, x_2, ..., x_n$ be the variables representing the numbers on the first, second, ..., nth ticket. The probability of drawing a ticket out of n tickets being in each case $1/n$, we have

$$E(x_i) = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + 3 \cdot \frac{1}{n} + ... + n \cdot \frac{1}{n} = \frac{1}{2}(n+1)$$

$\therefore$ expected value of the sum of the numbers on the tickets drawn

$$= E(x_1 + x_2 + ... + x_m) = E(x_1) + E(x_2) + ... + E(x_m)$$

$$= mE(x_i) = \frac{1}{2} m (n+1).$$

Example 26.33. *X is a continuous random variable with probability density function given by*

$$f(x) = kx \ (0 \le x < 2)$$
$$= 2k \ (2 \le x < 4)$$
$$= -kx + 6k \ (4 \le x < 6)$$

Find k and mean value of X. *(J.N.T.U., 2003)*

Solution. Since the total probability is unity

$\therefore$ $$\int_0^6 f(x)\, dx = 1$$

i.e., $$\int_0^2 kx\, dx + \int_2^4 2k dx + \int_4^6 (-kx + 6k)\, dx = 1$$

or $$k \left| x^2/2 \right|_0^2 + 2k \left| x \right|_2^4 + \left(-kx^2/2 + 6kx\right)_4^6 = 1$$

or $$2k + 4k + (-10k + 12k) = 1 \text{ i.e., } k = 1/8.$$

$$\text{Mean of } X = \int_0^6 x\, f(x)\, dx$$

$$= \int_0^2 kx^2\, dx + \int_2^4 2kx\, dx + \int_4^6 x(-kx + 6k) dx$$

$$= k \left| x^3/3 \right|_0^2 + 2k \left| x^2/2 \right|_2^4 + \left(-k \left| x^3/3 \right|_4^6 + 6k \left| x^2/2 \right|_4^6\right)$$

$$= k\,(8/3) + k\,(12) - k\,(152/3) + 3k\,(20) = \frac{1}{8}(24) = 3.$$

Example 26.34. *A variate X has the probability distribution*

x :	-3	6	9
$P(X = x)$:	$1/6$	$1/2$	$1/3$

Find E (X) and E (X²). Hence evaluate E (2X + 1)².

Solution. $$E(X) = -3 \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = 11/2.$$

$$E(X)^2 = 9 \times \frac{1}{6} + 36 \times \frac{1}{2} + 81 \times \frac{1}{3} = 93/2$$

$\therefore$ $$E(2X+1)^2 = E(4X^2 + 4X + 1) = 4E(X^2) + 4E(X) + 1$$
$$= 4\,(93/2) + 4\,(11/2) + 1 = 209.$$

Example 26.35. *The frequency distribution of a measurable characteristic varying between 0 and 2 is as under*

$$f(x) = x^3, \quad 0 \le x \le 1$$
$$= (2-x)^3, \quad 1 \le x \le 2.$$

Calculate the standard deviation and also the mean deviation about the mean.

Solution. Total frequency $N = \int_0^1 x^3\,dx + \int_1^2 (2-x)^3\,dx = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$$\therefore \quad \mu_1' \text{ (about the origin)} = \frac{1}{N}\left[\int_0^1 x \cdot x^3\,dx + \int_1^2 x(2-x)^3\,dx\right]$$

$$= 2\left\{\left|\frac{x^5}{5}\right|_0^1 + \left|-x \cdot \frac{(2-x)^4}{4}\right|_1^2 - \left|\frac{(2-x)^5}{20}\right|_1^2\right\} = 2\left(\frac{1}{5} + \frac{1}{4} + \frac{1}{20}\right) = 1$$

$$\mu_2' \text{ (about the origin)} = \frac{1}{N}\left[\int_0^1 x^2 \cdot x^3 dx + \int_1^2 x^2\,(2-x)^3\,dx\right]$$

$$= 2\left\{\left|\frac{x^6}{6}\right|_0^1 + \left|-x^2\,\frac{(2-x)^4}{4}\right|_1^2 + \frac{1}{2}\int_1^2 x\,(2-x)^4\,dx\right\}$$

$$= 2\left\{\frac{1}{6} + \frac{1}{4} + \frac{1}{2}\left[\frac{1}{5} + \frac{1}{30}\right]\right\} = \frac{16}{15}$$

Hence $\quad \sigma^2 = \mu_2 = \mu_2' - (\mu_1')^2 = \frac{1}{15}$

i.e., standard deviation $\sigma = \frac{1}{\sqrt{15}}$.

Mean deviation about the mean

$$= \frac{1}{N}\left\{\int_0^1 |x-1|\,x^3\,dx + \int_1^2 |x-1|\,(2-x)^3\,dx\right\}$$

$$= 2\left\{\int_0^1 (1-x)x^3\,dx + \int_1^2 (x-1)\,(2-x)^3\,dx\right\}$$

$$= 2\left\{\left(\frac{1}{4} - \frac{1}{5}\right) + \left(0 + \frac{1}{20}\right)\right\} = \frac{1}{5}.$$

26.11 MOMENT GENERATING FUNCTION

(1) *The moment generating function (m.g.f.) of the discrete probability distribution of the variate X about the value x = a is defined as the expected value of* $e^{t(x-a)}$ and is denoted by $M_a(t)$. Thus

$$M_a(t) = \Sigma p_i\, e^{t(x_i - a)} \qquad \ldots(1)$$

which is a function of the parameter t only.

Expanding the exponential in (1), we get

$$M_a(t) = \Sigma p_i + t\Sigma p_i\,(x_i - a) + \frac{t^2}{2!}\,\Sigma p_i(x_i - a)^2 + \ldots + \frac{t^r}{r!}\,\Sigma x_i\,(x_i - a)^r + \ldots$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\,\mu'_2 + \ldots + \frac{t^r}{r!}\mu'_r + \ldots \qquad \ldots(2)$$

where μ'_r is the moment of order r about a. Thus $M_a(t)$ generates moments and that is why it is called the moment generating function. From (2), we find

$$\mu'_r = \text{coefficient of } t^r/r! \text{ in the expansion of } M_a(t).$$

Otherwise differentiating (2) r times with respect to t and then putting $t = 0$, we get

$$\mu'_r = \left[\frac{d^r}{dt^r}\,M_a(t)\right]_{t=0} \qquad \ldots(3)$$

Thus the moment about any point $x = a$ can be found from (2) or more conveniently from the formula (3).

Rewriting (1) as

$$M_a(t) = e^{-at}\,\Sigma p_i\, e^{tx_i} \quad \text{or} \quad M_a(t) = e^{-at}\,M_0(t) \qquad \ldots(4)$$

Thus the m.g.f. about the point $a = e^{-at}$ *(m.g.f. about the origin).*

Obs. *The m.g.f. of the sum of two independent variables is the product of their m.g.fs.* ...(5)

(2) *If $f(x)$ is the density function of a continuous variate X, then the moment generating function of this continuous probability distribution about $x = a$ is given by*

$$M_a(t) = \int_{-\infty}^{\infty} e^{t(x-a)} f(x)\, dx.$$

Example 26.36. *Find the moment generating function of the exponential distribution*

$$f(x) = \frac{1}{c} e^{-x/c},\ 0 \le x \le \infty,\ c > 0.$$ *Hence find its mean and S.D.* (*Kurukshetra, 2009*)

Solution. The moment generating function about the origin is

$$M_0(t) = \int_0^{\infty} e^{tx} \cdot \frac{1}{c} e^{-x/c}\, dx = \frac{1}{c}\int_0^{\infty} e^{(t-1/c)x}\, dx \qquad \left[\because\ |t| < \frac{1}{c}\right]$$

$$= \frac{1}{c} \frac{\left| e^{(t-1/c)x} \right|_0^{\infty}}{|(t-1/c)|} = (1-ct)^{-1} = 1 + ct + c^2t^2 + c^3t^3 + \ldots$$

$\therefore$
$$\mu'_1 = \left[\frac{d}{dt} M_0(t)\right]_{t=0} = (c + 2c^2 t + 3c^3 t^2 + \ldots)_{t=0} = c$$

$$\mu'_2 = \left[\frac{d^2}{dt^2} M_0(t)\right]_{t=0} = 2c^2, \text{ and } \mu_2 = \mu'_2 - (\mu'_1)^2 = 2c^2 - c^2 = c^2.$$

Hence the mean is c and S.D. is also c.

26.12 PROBABILITY GENERATING FUNCTION

The probability generating function (p.g.f.) $P_x(t)$ for a random variable x which takes integral values 0, 1, 2, 3, ... only, is defined by

$$P_x(t) = p_0 + p_1 t + p_2 t^2 + \ldots = \sum_{n=0}^{\infty} p_n t^n = E(t^x)$$

The coefficient of t^n in the expansion of $P(t)$ in powers of t gives $P(t)_{x=n}$.

$$\frac{\partial P}{\partial t} = \sum_{n=0}^{\infty} n p_n t^{n-1} \quad \text{or} \quad \left(\frac{\partial P}{\partial t}\right)_{t=1} = \Sigma\, n p_n = \mu_1'$$

$$\frac{\partial^2 P}{\partial t^2} = \sum_{n=0}^{\infty} n(n-1)\, p_n t^{n-2} \quad \text{or} \quad \left(\frac{\partial^2 P}{\partial t^2}\right)_{t=1} = \Sigma\, n(n-1)\, p_n = \mu_2' - \mu_1'$$

$$= \mu_2 + \mu_1'^2 - \mu_1' \text{ and so on}$$

Also $$\left(\frac{\partial^k P}{\partial t^n}\right)_{t=0} = n!\, p_n,\ k = 1, 2, \ldots n.$$

For integral valued variates, we have

$$P_x(e^t) = E(e^{tx}) = m.g.f. \text{ for } x.$$

Obs. The p.g.f. of the sum of two independent random variables is the product of their p.g.f.'s.

Example 26.37. *If x be a random variable with probability generating function $P_x(t)$, find the probability generating function of*

(*i*) $x + 2$ (*ii*) $2x$.

Solution. We have $P_x(t) = \sum_{k=0}^{\infty} p_k\, t^k$

(*i*) Probability generating function of $x + 2 = \sum_{k=0}^{\infty} p_k\, t^{k+2} = t^2 \sum_{k=0}^{\infty} p_k\, t^k = t^2 P_x(t).$

(*ii*) Probability generating function of $2x = \sum_{k=0}^{\infty} p_k\, t^{2k} = \sum_{k=0}^{\infty} p_k (t^2)^k = P(t^2)$.

PROBLEMS 26.4

1. A random variable x has the following probability function :

Values of x :	-2	-1	0	1	2	3
$p(x)$:	0.1	k	0.2	$2k$	0.3	k

Find the value of k and calculate mean and variance. (*S.V.T.U., 2007 ; V.T.U., 2004 ; Madras, 2003*)

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

x :	8	12	16	20	24
$p(x)$:	1/8	1/6	3/8	1/4	1/12

3. Obtain the distribution function of the total number of heads occurring in three tosses of an unbiased coin.
4. Show that for any discrete distribution $\beta_2 \geq 1$.
5. From an urn containing 3 red and 2 white balls, a man is to draw 2 balls at random without replacement, being promised Rs. 20 for each red ball he draws and Rs. 10 for each white one. Find his expectation.
6. Four coins are tossed. What is the expectation of the number of heads ?
7. The diameter of an electric cable is assumed to be a continuous variate with p.d.f. $f(x) = 6x(1-x)$, $0 \leq x \leq 1$. Verify that the above is a p.d.f. Also find the mean and variance.
8. A random variable gives measurements X between 0 and 1 with a probability function

$$f(x) = 12x^3 - 21x^2 + 10x, \quad 0 \leq x \leq 1$$
$$= 0$$

(*i*) Find $P\left(X \leq \frac{1}{2}\right)$ and $P\left(X > \frac{1}{2}\right)$

(*ii*) Find a number k such that $P(X \leq k) = \frac{1}{2}$ (*J.N.T.U., 2003*)

9. The power reflected by an aircraft that is received by a radar can be described by an exponential random variable X.

The probability density of X is given by $f(x) = \begin{cases} \frac{1}{x_0} e^{-x/x_0}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

where x_0 is the average power received by the radar.

(*i*) What is the probability that the radar will receive power larger than the power received on the average ? (*ii*) What is the probability that the radar will receive power less than the power received on the average ?

(*Mumbai, 2006*)

10. A function is defined as follows :

$$f(x) = 0, \quad x < 2$$
$$= \frac{1}{18}(2x+3), \quad 2 \leq x \leq 4$$
$$= 0, \quad x > 4.$$

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

11. A continuous distribution of a variable x in the range $(-3, 3)$ is defined as

$$f(x) = \frac{1}{16}(3+x)^2, \quad -3 \leq x < -1$$
$$= \frac{1}{16}(2-6x^2), \quad -1 \leq x < 1$$
$$= \frac{1}{16}(3-x)^2, \quad 1 \leq x \leq 3.$$

Verify that the area under the curve is unity. Show that the mean is zero. *(Kurukshetra, 2005)*

12. The frequency function of a continuous random variable is given by

$$f(x) = y_0 x (2 - x),\ 0 \le x \le 2.$$

Find the value of y_0, mean and variance of x. *(Kerala, 2005 ; J.N.T.U., 2003)*

13. The probability density $p(x)$ of a continuous random variable is given by

$$p(x) = y_0 e^{-|x|},\ -\infty < x < \infty.$$

Prove that $y_0 = 1/2$. Find the mean and variance of the distribution.

(S.V.T.U., 2008 ; Kurukshetra, 2007 ; V.T.U., 2004)

14. If $f(x) = \begin{cases} \frac{1}{2}(x+1), & -1 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$

represents the density of a random variable X, find $E(X)$ and Var (X).

15. A function is defined as under :

$$f(x) = 1/k,\ x_1 \le x \le x_2$$
$$= 0, \text{ elsewhere.}$$

Find the cumulative distribution of the variate x when k satisfies the requirements for $f(x)$ to be a density function.

26.13 REPEATED TRIALS

We know that the probability of getting a head or a tail on tossing a coin is $\frac{1}{2}$. If the coin is tossed thrice, the probability of getting one head and two tails can be combined as $H—T—T$, $T—H—T$, $T—T—H$. The probability of each one of these being $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$, *i.e.*, $\left(\frac{1}{2}\right)^3$, their total probability shall be $3(1/2)^3$.

Similarly if a trial is repeated n times and if p is the probability of a success and q that of a failure, then the probability of r successes and $n - r$ failures is given by $p^r q^{n-r}$.

But these r successes and $n - r$ failures can occur in any of the ${}^nC_r$ ways in each of which the probability is same.

Thus the probability of r successes is ${}^nC_r\, p^r q^{n-r}$.

Cor. *The probabilities of at least r successes in n trials*

= the sum of the probabilities of $r, r+1, \ldots, n$ successes

$$= {}^nC_r\, p^r q^{n-r} + {}^nC_{r+1}\, p^{r+1} q^{n-r-1} + \ldots + {}^nC_n\, p^n.$$

26.14 (1) BINOMIAL DISTRIBUTION*

It 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.

If we perform a series of independent trials such that for each trial p is the probability of a success and q that of a failure, then the probability of r successes in a series of n trials is given by ${}^nC_r\, p^r q^{n-r}$, where r takes any integral value from 0 to n. The probabilities of 0, 1, 2, ... r, ..., n successes are, therefore, given by

$$q^n,\ {}^nC_1 pq^{n-1},\ {}^nC_2 p^2 q^{n-2},\ \ldots,\ {}^nC_r\, p^r q^{n-r},\ \ldots,\ p^n.$$

The probability of the number of successes so obtained is called *the* **binomial distribution** for the simple reason that the probabilities are the successive terms in the expansion of the binomial $(q + p)^n$.

∴ the sum of the probabilities

$$= q^n + {}^nC_1 pq^{n-1} + {}^nC_2 p^2 q^{n-2} + \ldots + p^n = (q+p)^n = 1.$$

(2) Constants of the binomial distribution. The moment generating function about the origin is

$$M_0(t) = E(e^{tx}) = \Sigma\, {}^nC_x\, p^x q^{n-x} e^{tx} \qquad \text{[By (1) § 26.11]}$$
$$= \Sigma\, {}^nC_x (pe^t)^x q^{n-x} = (q + pe^t)^n$$

* It was discovered by a Swiss mathematician *Jacob Bernoulli* and was published posthumously in 1713.

Differentiating with respect to t and putting $t = 0$ and using (3) § 26.11, we get the mean

$$\mu'_1 = np.$$

Since $M_a(t) = e^{-at} M_0(t)$, the m.g.f. of the binomial distribution about its mean $(m) = np$, is given by

$$M_m(t) = e^{-npt} (q + pe^t)^n = (qe^{-pt} + pe^{qt})^n$$

$$= \left(1 + pq \frac{t^2}{2!} + pq(q^2 - p^2)\frac{t^3}{3!} + pq(q^3 + p^3)\frac{t^4}{4!} + \cdots\right)^n$$

or
$$1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \mu_3 \frac{t^3}{3!} + \mu_4 \frac{t^4}{4!} + \cdots$$

$$= 1 + npq \frac{t^2}{2!} + npq(q-p)\frac{t^3}{3!} + npq\,[1 + 3(n-2)pq]\frac{t^4}{4!} + \ldots$$

Equating the coefficients of like powers of t on either side, we have

$$\mu_2 = npq,\ \mu_3 = npq(q-p),\ \mu_4 = npq\,[1 + 3(n-2)pq].$$

Also
$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

Thus *mean* $= np$, *standard deviation* $= \sqrt{(npq)}$.

skewness $= (1-2p)/\sqrt{(npq)}$, *kurtosis* $= \beta_2$.

Obs. The skewness is positive for $p < \frac{1}{2}$ and negative for $p > \frac{1}{2}$. When $p = \frac{1}{2}$, the skewness is zero, *i.e.*, the probability curve of the binomial distribution will be symmetrical (bell-shaped).

As n the number of trials increase indefinitely, $\beta_1 \to 0$, and $\beta_2 \to 3$.

(3) Binomial frequency distribution. If n independent trials constitute one experiment and this experiment be repeated N times, then the frequency of r successes is $N\,{}^nC_r\, p^r q^{n-r}$. The possible number of successes together with these expected frequencies constitute the *binomial frequency distribution.*

(4) Applications of Binomial distribution. This distribution is applied to problems concerning :
(*i*) Number of defectives in a sample from production line,

(*ii*) Estimation of reliability of systems,

(*iii*) Number of rounds fired from a gun hitting a target,

(*iii*) Radar detection.

Example 26.38. *The probability that a pen manufactured by a company will be defective is 1/10. If 12 such pens are manufactured, find the probability that*

(*a*) *exactly two will be defective.* (*b*) *at least two will be defective.*

(*c*) *none will be defective.* (*V.T.U., 2004 ; Burdwan, 2003*)

Solution. The probability of a defective pen is 1/10 = 0.1

∴ The probability of a non-defective pen is 1 – 0.1 = 0.9

(*a*) *The probability that exactly two will be defective*

$$= {}^{12}C_2\,(0.1)^2\,(0.9)^{10} = 0.2301$$

(*b*) *The probability that at least two will be defective*

$$= 1 - (\text{prob. that either none or one is non-defective})$$
$$= 1 - [{}^{12}C_0(0.9)^{12} + {}^{12}C_1(0.1)\,(0.9)^{11}] = 0.3412$$

(*c*) *The probability that none will be defective*

$$= {}^{12}C_{12}(0.9)^{12} = 0.2833.$$

Example 26.39. *In 256 sets of 12 tosses of a coin, in how many cases one can expect 8 heads and 4 tails.* (*J.N.T.U., 2003*)

Solution. $P(\text{head}) = \frac{1}{2}$ and $P(\text{tail}) = \frac{1}{2}$

By binomial distribution, probability of 8 heads and 4 tails in 12 trials is

$$P(X=8) = {}^{12}C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^4 = \frac{12!}{8!\,4!} \cdot \frac{1}{2^{12}} = \frac{495}{4096}$$

$\therefore$ the expected number of such cases in 256 sets

$$= 256 \times P(X=8) = 256\ \frac{495}{4096} = 30.9 = 31 \text{ (say)}.$$

Example 26.40. *In sampling a large number of parts manufactured by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples, how many would be expected to contain at least 3 defective parts.* *(V.T.U., 2004)*

Solution. Mean number of defectives $= 2 = np = 20p$.

$\therefore$ The probability of a defective part is $p = 2/20 = 0.1$.

and the probability of a non-defective part = 0.9

$\therefore$ The probability of at least three defectives in a sample of 20.

$$= 1 - (\text{prob. that either none, or one, or two are non-defective parts})$$
$$= 1 - [{}^{20}C_0(0.9)^{20} + {}^{20}C_1(0.1)(0.9)^{19} + {}^{20}C_2(0.1)^2(0.9)^{18}]$$
$$= 1 - (0.9)^{18} \times 4.51 = 0.323.$$

Thus *the number of samples having at least three defective parts* out of 1000 samples

$$= 1000 \times 0.323 = 323.$$

Example 26.41. *The following data are the number of seeds germinating out of 10 on damp filter paper for 80 sets of seeds. Fit a binomial distribution to these data :*

x :	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
f :	*6*	*20*	*28*	*12*	*8*	*6*	*0*	*0*	*0*	*0*	*0*

Solution. Here $n = 10$ and $N = \Sigma f_i = 80$

$$\therefore \qquad \text{mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{20+56+36+32+30}{80} = \frac{174}{80} = 2.175$$

Now the mean of a binomial distribution $= np$

i.e., $\quad np = 10p = 2.175 \quad \therefore \quad p = 0.2175,\ q = 1 - p = 0.7825$

Hence the binomial distribution to be fitted is

$$N(q+p)^n = 80\,(0.7825 + 0.2175)^{10}$$
$$= 80.{}^{10}C_0\,(0.7825)^{10} + 80.{}^{10}C_1\,(0.7825)^9\,(0.2175)^1 + {}^{10}C_2\,(0.7825)^8\,(0.2175)^2 + \ldots + {}^{80}C_9\,(0.7825)^1\,(0.2175)^9 + {}^{80}C_{10}\,(0.2175)^{10}$$
$$= 6.885 + 19.13 + 23.94 + \ldots + 0.0007 + 0.00002$$

$\therefore$ the successive terms in the expansion give the expected or theoretical frequencies which are

x :	0	1	2	3	4	5	6	7	8	9	10
f :	6.9	19.1	24.0	17.8	8.6	2.9	0.7	0.1	0	0	0

PROBLEMS 26.5

1. Determine the binomial distribution for which mean = 2 (variance) and mean + variance = 3. Also find $P(X \le 3)$. *(Kerala, 2005)*
2. An ordinary six-faced die is thrown four times. What are the probabilities of obtaining 4, 3, 2, 1 and 0 faces ?
3. If the chance that one of the ten telephone lines is busy at an instant is 0.2.
   (*a*) What is the chance that 5 of the lines are busy ?
   (*b*) What is the most probable number of busy lines and what is the probability of this number ?
   (*c*) What is the probability that all the lines are busy ? *(V.T.U., 2002 S)*
4. If the probability that a new-born child is a male is 0.6, find the probability that in a family of 5 children there are exactly 3 boys. *(Kurukshetra, 2005)*

5. If on an average 1 vessel in every 10 is wrecked, find the probability that out of 5 vessels expected to arrive, at least 4 will arrive safely. *(P.T.U., 2005)*
6. The probability that a bomb dropped from a plane will strike the target is 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.
7. A sortie of 20 aeroplanes is sent on an operational flight. The chances that an aeroplane fails to return is 5%. Find the probability that (*i*) one plane does not return (*ii*) at the most 5 planes do not return, and (*iii*) what is the most probable number of returns ? *(Hissar, 2007)*
8. The probability that an entering student will graduate is 0.4. Determine the probability that out of 5 students (*a*) none (*b*) one and (*c*) at least one will graduate.
9. Out of 800 families with 5 children each, how many would you expect to have (*a*) 3 boys, (*b*) 5 girls, (*c*) either 2 or 3 boys ? Assume equal probabilities for boys and girls. *(V.T.U., 2004)*
10. If 10 per cent of the rivets produced by a machine are defective, find the probability that out of 5 rivets chosen at random (*i*) none will be defective, (*ii*) one will be defective, and (*iii*) at least two will be defective.
11. 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. *(V.T.U., 2003 S)*
12. A product is 0.5% defective and is packed in cartons of 100. What percentage contains not more than 3 defectives ?
13. If in a lot of 500 solenoids 25 are defective, find the probability of 0, 1, 2, 3 defective solenoids in a random sample of 20 solenoids.
14. 500 articles were selected at random out of a batch containing 10,000 articles, and 30 were found to be defective. How many defectives articles would you reasonably expect to have in the whole batch ? *(J.N.T.U., 2003)*
15. 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

(Bhopal, 2006)

16. Fit a binomial distribution to the following frequency distribution :

x:	0	1	2	3	4	5	6
f:	13	25	52	58	32	16	4

(Kurukshetra, 2009 ; S.V.T.U., 2007)

26.15 (1) POISSON DISTRIBUTION*

It is a distribution related to the probabilities of events which are extremely rare, but which have a large number of independent opportunities for occurrence. The number of persons born blind per year in a large city and the number of deaths by horse kick in an army corps are some of the phenomena, in which this law is followed.

This distribution can be derived as a limiting case of the binomial distribution by making n very large and p very small, keeping np fixed (= m, say).

The probability of r successes in a binomial-distribution is

$$P(r) = {}^nC_r\, p^r q^{n-r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}\, p^r q^{n-r}$$

$$= \frac{np(np-p)(np-2p)\cdots(np-\overline{r-1}p)}{r!}\,(1-p)^{n-r}$$

As $n \to \infty$, $p \to 0$ $(np = m)$, we have

$$P(r) = \frac{m^r}{r!} \operatorname*{Lt}_{n\to\infty} \frac{(1-m/n)^n}{(1-m/n)^r} = \frac{m^r}{r!}\, e^{-m}$$

so that the probabilities of 0, 1, 2..., r,... successes in a Poisson distribution are given by

$$e^{-m},\ me^{-m},\ \frac{m^2e^{-m}}{2!},\cdots,\frac{m^re^{-m}}{r!},\cdots$$

The sum of these probabilities is unity as it should be.

* It was discovered by a French mathematician S.D. Poisson in 1837.

(2) Constants of the Poisson distribution. These constants can easily be derived from the corresponding constants of the binomial distribution simply by making $n \to \infty$, $p \to 0$, $(q \to 1)$ and noting that $np = m$

$$Mean = \text{Lt}\,(np) = m$$

$$\mu_2 = \text{Lt}\,(npq) = m\ \text{Lt}\,(q) = m$$

$\therefore$ *Standard deviation* $= \sqrt{m}$

Also $\mu_3 = m$, $\mu_4 = m + 3m^2$

$\therefore$ *Skewness* $(= \sqrt{\beta_1}) = 1/m$, *Kurtosis* $(= \beta_2) = 3 + 1/m$.

Since μ_3 is positive, Poisson distribution is positively skewed and since $\beta_2 > 3$, it is *Leptokurtic.*

(3) Applications of Poisson distribution. This distribution is applied to problems concerning : (*i*) Arrival pattern of 'defective vehicles in a workshop', 'patients in a hospital' or 'telephone calls'.

(*ii*) Demand pattern for certain spare parts.

(*iii*) Number of fragments from a shell hitting a target.

(*iv*) Spatial distribution of bomb hits.

Example 26.42. *If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 2,000 individuals more than two will get a bad reaction.* (V.T.U., 2008 ; Kottayam, 2005)

Solution. It follows a Poisson distribution as the probability of occurrence is very small.

Mean $m = np = 2000(0.001) = 2$

Probability that more than 2 will get a bad reaction

= 1 – [prob. that no one gets a bad reaction + prob. that one gets a bad reaction + prob. that two get bad reaction]

$$= 1 - \left[e^{-m} + \frac{m^1 e^{-m}}{1!} + \frac{m^2 e^{-m}}{2!}\right] = 1 - \left[\frac{1}{e^2} + \frac{2}{e^2} + \frac{2}{e^2}\right] \quad [\because \ m = 2]$$

$$= 1 - \frac{5}{e^2} = 0.32. \quad [\because \ e = 2.718]$$

Example 26.43. *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, use Poisson distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 10,000 packets.* (Kurukshetra, 2009 S; Madras, 2006 ; V.T.U., 2004)

Solution. We know that $m = np = 10 \times 0.002 = 0.02$

$$e^{-0.02} = 1 - 0.02 + \frac{(0.02)^2}{2!} - \ldots = 0.9802 \text{ approximately}$$

Probability of no defective blade is $e^{-m} = e^{-0.02} = 0.9802$

$\therefore$ no. of packets containing no defective blade is

$$10{,}000 \times 0.9802 = 9802$$

Similarly the number of packets containing one defective blade $= 10{,}000 \times me^{-m}$

$$= 10{,}000 \times (0.02) \times 0.9802 = 196$$

Finally the number of packets containing two defective blades

$$= 10{,}000 \times \frac{m^2 e^{-m}}{2!} = 10{,}000 \times \frac{(0.02)^2}{2!} \times 0.9802 = 2 \text{ approximately.}$$

Example 26.44. *Fit a Poisson distribution to the set of observations :*

x :	0	1	2	3	4
f :	122	60	15	2	1

(Bhopal, 2007 S ; V.T.U., 2004 ; U.P.T.U., 2003)

Solution. Mean $= \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{60 + 36 + 6 + 4}{200} = 0.5.$

∴ mean of Poisson distribution *i.e.*, $m = 0.5$.

Hence the theoretical frequency for r successes is

$$\frac{Ne^{-m}(m)^r}{r!} = \frac{200e^{-0.5}(.5)^r}{r!} \quad \text{where } r = 0, 1, 2, 3, 4$$

∴ the theoretical frequencies are

x:	0	1	2	3	4
f:	121	61	15	2	0

($\because$ $e^{-.5} = 0.61$)

PROBLEMS 26.6

1. If a random variable has a Poisson distribution such that $P(1) = P(2)$, find
   (*i*) mean of the distribution. (*ii*) $P(4)$. (*V.T.U., 2003*)
2. X is a Poisson variable and it is found that the probability that $X = 2$ is two-thirds of the probability that $X = 1$. Find the probability that $X = 0$ and the probability that $X = 3$. What is the probability that X exceeds 3 ?
3. For Poisson distribution, prove that $m\,\mu_2\,\gamma_1\,\gamma_2 = 1$, where symbols have their usual meanings. (*S.V.T.U., 2008*)
4. A certain screw making machine produces on average of 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws. (*Kurukshetra, 2006*)
5. A manufacturer knows that the condensers he makes contain on the average 1 % defectives. He packs them in boxes of 100. What is the probability that a box picked at random will contain 3 or more faulty condensers ?
6. 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 (*i*) on which there is no demand, (*ii*) on which demand is refused. ($e^{-1.5} = 0.2231$). (*Bhopal, 2008 S ; J.N.T.U., 2003*)
7. The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is probability that in a group of 7, five or more will suffer from it ?
8. The frequency of accidents per shift in a factory is as shown in the following table :

Accidents per shift :	0	1	2	3	4
Frequency :	180	92	24	3	1

   Calculate the mean number of accidents per shift and the corresponding Poisson distribution and compare with actual observations.
9. A source of liquid is known to contain bacteria with the mean number of bacteria per cubic centimetre equal to 3. Ten 1 c.c., test-tubes are filled with the liquid. Assuming that Poisson distribution is applicable, calculate the probability that all the test-tubes will show growth *i.e.*, contain atleast 1 bacterium each.
10. Find the expectation of the function $\phi(x) = xe^{-x}$ in a Poisson distribution. (*V.T.U., 2003*)
   [**Hint** : If m be the mean of the Poisson distribution, then expectation of
   $$\phi(x) = \sum_{x=0}^{\infty} \frac{\phi(x)\cdot m^x e^{-m}}{x!} = m \exp.\, m\,(e^{-1} - m - 1)$$
11. Fit a Poisson distribution to the following :

x:	0	1	2	3	4
f:	46	38	22	9	1

   (*Kurukshetra, 2009 ; Bhopal, 2008 ; V.T.U., 2003 S*)
12. 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

   (*S.V.T.U., 2007*)

26.16 (1) NORMAL DISTRIBUTION*

Now we consider a continuous distribution of fundamental importance, namely the normal distribution. *Any quantity whose variation depends on random causes is distributed according to the normal law.* Its importance lies in the fact that a large number of distributions approximate to the normal distribution.

---

* In 1924, *Karl Pearson* found this distribution which *Abraham De Moivre* had discovered as early as 1733. See footnote p. 843 and 647.

Let us define a variate $z = \dfrac{x - np}{\sqrt{(npq)}}$ ...(1)

where x is a binomial variate with mean np and S.D. $\sqrt{(npq)}$ so that z is a variate with mean zero and variance unity. In the limit as n tends to infinity, the distribution of z becomes a continuous distribution extending from $-\infty$ to ∞.

It can be shown that *the limiting form of the binomial distribution (1) for large values of n when neither p nor q is very small, is the normal distribution.* The normal curve is of the form

$$y = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} \quad ...(2)$$

where μ and σ are the mean and standard deviation respectively.

(2) Properties of the normal distribution

I. The normal curve (2) is bell-shaped and is symmetrical about its mean. It is unimodal with ordinates decreasing rapidly on both sides of the mean (Fig. 26.3). The maximum ordinate is $1/\sigma\sqrt{(2\pi)}$, found by putting $x = \mu$ in (2).

As it is symmetrical, its mean, median and mode are the same. Its points of inflexion (found by putting $d^2y/dx^2 = 0$ and verifying that at these points $d^3y/dx^3 \neq 0$) are given by $x = \mu \pm \sigma$, *i.e.*, these points are equidistant from the mean on either side.

II. Mean deviation from the mean μ

$$= \int_{-\infty}^{\infty} |x-\mu| \cdot \frac{1}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} dx \qquad [\text{Put } z = (x-\mu)/\sigma]$$

$$= \frac{\sigma}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} |z| e^{-z^2/2} dz$$

$$= \frac{\sigma}{\sqrt{(2\pi)}} \left[\int_{-\infty}^{0} -ze^{-z^2/2}dz + \int_{0}^{\infty} ze^{-z^2/2}dz \right] = \frac{2\sigma}{\sqrt{(2\pi)}} \int_0^{\infty} ze^{-z^2/2} dz$$

$$= \frac{2\sigma}{\sqrt{(2\pi)}} \left| -e^{-z^2/2} \right|_0^{\infty} = -\sqrt{\left(\frac{2}{\pi}\right)} \sigma(0-1) = 0.7979\,\sigma \approx (4/5)\,\sigma$$

III. Moments about the mean

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} \frac{1}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} dx$$

$$= \frac{\sigma^{2n+1}}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} z^{2n+1} e^{-z^2/2} dz \text{ where } z = (x-\mu)/\sigma$$

= 0, since the integral is an odd function.

Thus all odd order moments about the mean vanish.

$$\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} \frac{1}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} dx$$

$$= \frac{\sigma^{2n}}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} z^{2n-1} e^{-z^2/2} . zdz \qquad [\text{Integrate by parts}]$$

$$= \frac{\sigma^{2n}}{\sqrt{(2\pi)}} \left[\left| -z^{2n-1} e^{-z^2/2} \right|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} (2n-1) z^{2n-2} e^{-z^2/2} dz \right]$$

$$= \frac{\sigma^{2n}}{\sqrt{(2\pi)}} (0-0) + (2n-1)\,\sigma^2 \mu_{2n-2}$$

Repeated application of this reduction formula, gives

$$\mu_{2n} = (2n-1)(2n-3) \ldots 3 \,.\, 1\sigma^{2n}$$

In particular, $\mu_2 = \sigma^2$, $\mu_4 = 3\sigma^4$.

Hence $\quad \beta_1 = \dfrac{\mu_3^{\,2}}{\mu_2^{\,3}} = 0$ and $\beta_2 = \dfrac{\mu_4}{\mu_2^{\,2}} = 3$

i.e., the coefficient of skewness is zero (*i.e.* the curve is symmetrical) and the Kurtosis is 3. This is the basis for the choice of the value 3 in the definitions of platykurtic and leptokurtic (page 844).

IV. The probability of x lying between x_1 and x_2 is given by the area under the normal curve from x_1 to x_2, *i.e.*, $P\ (x_1 \le x \le x_2)$

$$= \frac{1}{\sigma\sqrt{(2\pi)}} \int_{x_1}^{x_2} e^{-(x-\mu)^2/2\sigma^2}\, dx$$

$$= \frac{1}{\sqrt{(2\pi)}} \int_{z_1}^{z_2} e^{-z^2/2}\, dz \text{ where } z = (x-\mu)/\sigma,\ dz = dx/\sigma \text{ and } z_1 = (x_1-\mu)/\sigma,\ z_2 = (x_2-\mu)/\sigma.$$

$$= \frac{1}{\sqrt{(2\pi)}} \left\{ \int_0^{z_2} e^{-z^2/2}\, dz - \int_0^{z_1} e^{-z^2/2}\, dz \right\} = P_2(z) - P_1(z)$$

The values of each of the above integrals can be found from the table III–Appendix 2, which gives the values of

$$P(z) = \frac{1}{\sqrt{(2\pi)}} \int_0^{z} e^{-z^2/2}\, dz$$

for various values of z. This integral is called the *probability integral* or the *error function* due to its use in the theory of sampling and the theory of errors.

Using this table, we see that the area under the normal curve from $z = 0$ to $z = 1$, *i.e.* from $x = \mu$ to $\mu + \sigma$ is 0.3413.

∴ (*i*) The area under the normal curve between the ordinates $x = \mu - \sigma$ and $x = \mu + \sigma$ is 0.6826, ~ *68% nearly. Thus approximately 2/3 of the values lie within these limits.*

(*ii*) The area under the normal curve between $x = \mu - 2\sigma$ and $x = \mu + 2\sigma$ is 0.9544 ~ 95.5%, which implies that *about $4\frac{1}{2}$ % of the values lie outside these limits.*

(*ii*) *99.73% of the values lie between $x = \mu - 3\sigma$ and $x = \mu + 3\sigma$ i.e., only a quarter % of the whole lies outside these limits.*

(*iv*) *95% of the values lie between $x = \mu - 1.96\sigma$ and $x = \mu + 1.96\,\sigma$ i.e., only 5% of the values lie outside these limits.*

(*v*) *99% of the values lie between $x = \mu - 2.58\sigma$ and $x = \mu + 2.58\sigma$ i.e., only 1% of the values lie outside these limits.*

(*vi*) 99.9% of the values lie between $x = \mu - 3.29\sigma$ and $x = \mu + 3.29\sigma$.

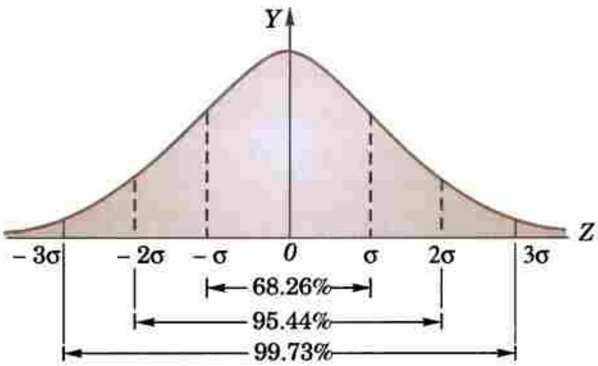

Fig. 26.4

In other words, a value that deviates more than σ from μ occurs about once in 3 trials. A value that deviates more than 2σ or 3σ from μ occurs about once in 20 or 400 trials. Almost all values lie within 3σ of the mean.

The shape of the standardised normal curve is

$$y = \frac{1}{\sqrt{(2\pi)}} e^{-z^2/2} \text{ where } z = (x-\mu)/\sigma \qquad ...(3)$$

and the respective areas are shown in Fig. 26.4. 'z' is called a *normal variate*.

(3) Normal frequency distribution. We can fit a normal curve to any distribution. If N be the total frequency, μ the mean and σ the standard deviation of the given distribution then the curve

$$y = \frac{N}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} \qquad ...(4)$$

will fit the given distribution as best as the data will permit. The frequency of the variate between x_1 and x_2 as given by the fitted curve, will be the area under (1) from x_1 to x_2.

(4) Applications of normal distribution. This distribution is applied to problems concerning :

(*i*) Calculation of errors made by chance in experimental measurements.

(*ii*) Computation of hit probability of a shot.

(*iii*) Statistical inference in almost every branch of science.

26.17 PROBABLE ERROR

Any lot of articles manufactured to certain specifications is subject to small errors. In fact, measurement of any physical quantity shows slight error. In general, these errors of manufacture or experiment are of random nature and therefore, follow a normal distribution. While quoting a specification of an experimental result, we usually mention the *probable error* (λ). *It is such that the probability of an error falling within the limits* $\mu - \lambda$ *and* $\mu + \lambda$ *is exactly equal to the chance of an error falling outside these limits, i.e.* the chance of an error lying within $\mu - \lambda$ and $\mu + \lambda$ is $\frac{1}{2}$.

$$\therefore \quad \frac{1}{\sigma\sqrt{(2\pi)}} \int_{\mu-\lambda}^{\mu+\lambda} e^{-(x-\mu)^2/2\sigma^2}\, dx = \frac{1}{2}$$

or

$$\frac{1}{\sqrt{(2\pi)}} \int_0^{\lambda/\sigma} e^{-z^2/2} dz = \frac{1}{4} \qquad \left[z = \frac{x-\mu}{\sigma}\right]$$

The table V, (Appendix 2) gives $\lambda/\sigma = 0.6745$

Hence the probable error $\lambda = 0.6745\sigma \sim \frac{2}{3}\sigma$.

Obs. Quartile deviation $= \frac{1}{2}(Q_3 - Q_1) \sim \frac{2}{3}\sigma$; Mean deviation $\approx \frac{4}{5}\sigma$ [p. 839]

$\therefore$ *Q.D. : M.D. : S.D.* = 10 : 12 : 15. *(Madras, 2003)*

Example 26.45. *X is a normal variate with mean 30 and S.D. 5, find the probabilities that (i) $26 \le X \le 40$, (ii) $X \ge 45$ and (iii) $|X - 30| > 5$.* *(J.N.T.U., 2005)*

Solution. We have $\mu = 30$ and $\sigma = 5$

$$\therefore \quad z = \frac{X-\mu}{\sigma} = \frac{X-30}{5}$$

(*i*) When $X = 26, z = -0.8$; when $X = 40, z = 2$

$$\begin{aligned}
\therefore \quad P(26 \le X \le 40) &= P(-0.8 \le z \le 2) \\
&= P(-0.8 \le z \le 0) + P(0 \le z \le 2) \quad \text{[Using Table III]} \\
&= P(0 \le z \le 0.8) + 0.4772 \\
&= 0.2881 + 0.4772 = 0.7653
\end{aligned}$$

(*ii*) When $X = 45, z = 3$

$$\begin{aligned}
\therefore \quad P(X \ge 45) &= P(z \ge 3) = 0.5 - P(0 \le z \le 3) \\
&= 0.5 - 0.4986 = 0.0014
\end{aligned}$$

(*iii*)

$$\begin{aligned}
P[|X-30| \le 5] &= P[25 \le X \le 35] \\
&= P(-1 \le z \le 1) = 2P(0 \le z \le 1) \\
&= 2 \times 0.3413 = 0.6826 \\
\therefore \quad P[|X-30| > 5] &= 1 - P[|X-30| \le 5] \\
&= 1 - 0.6826 = 0.3174.
\end{aligned}$$

Example 26.46. *A certain number of articles manufactured in one batch were classified into three categories according to a particular characteristic, being less than 50, between 50 and 60 and greater than 60. If this characteristic is known to be normally distributed, determine the mean and standard deviation for this batch if 60%, 35% and 5% were found in these categories.*

Solution. Let μ be the mean (at $z = 0$) and σ the standard deviation of the normal curve (Fig. 26.5).

Now 60% of the articles have the characteristic below 50, 35% between 50 and 60 and only 5% greater than 60.

Let the area to the left of the ordinate *PQ* be 60% and that between the ordinates *PQ* and *ST* be 35% so that the areas to the left of *PQ* ($z = z_1$) and *ST* ($z = z_2$) are 0.6 and 0.95 respectively, *i.e.*, the area *OPQR* = 0.6 – 0.5 = 0.1 and the area *OSTR* = 0.45.

$\therefore$ area corresponding to $z_1 \left(= \dfrac{50-\mu}{\sigma}\right) = 0.1$

and that corresponding to $z_2 \left(= \dfrac{60-\mu}{\sigma}\right) = 0.45$

From the table III, we have

$$(50-\mu)/\sigma = 0.2533 \quad \text{and} \quad (60-\mu)/\sigma = 1.645$$

whence $\sigma = 7.543$ and $\mu = 48.092.$

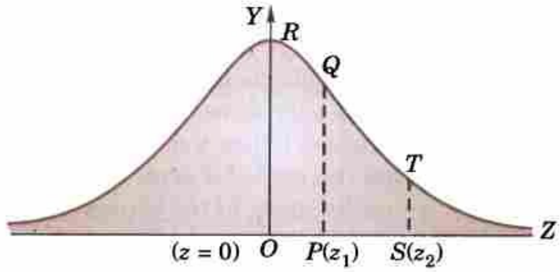

Fig. 26.5

Example 26.47. *In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and S.D. of the distribution.* *(V.T.U., 2009 ; S.V.T.U., 2008 ; Kurukshetra, 2007 S)*

Solution. Let $\bar{x}$ be the mean and σ the S.D. 31% of the items are under 45 means area to the left of the ordinate $x = 45$. (Fig. 26.6)

When $x = 45$, let $z = z_1$ so that $z_1 = \dfrac{45-\bar{x}}{\sigma}$ ...(*i*)

$$\therefore \quad \int_{-\infty}^{z_1} \phi(z)\,dz = 0.31 \quad \text{or} \quad \int_{-\infty}^{0} \phi(z)\,dz - \int_{z_1}^{0} \phi(z)\,dz = 0.31$$

Hence $$\int_{z_1}^{0} \phi(z)\,dz = \int_{-\infty}^{0} \phi(z)\,dz - 0.31 = 0.5 - 0.31 = 0.19$$

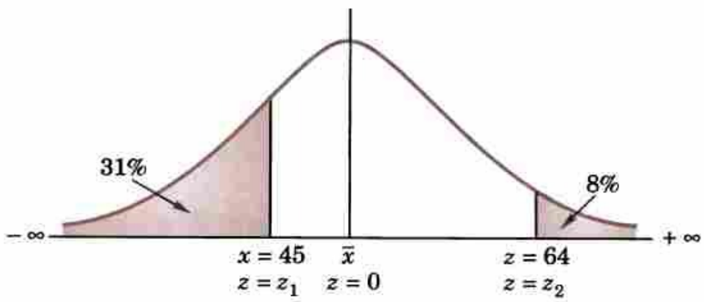

Fig. 26.6

From table III, $z_1 = -0.5$ ...(*ii*)

When $x = 64$, let $z = z_2$ so that $z_2 = (64 - \bar{x})/\sigma$ ...(*iii*)

$$\therefore \quad \int_{z_2}^{\infty} \phi(z)\,dz = 0.08 \quad \text{or} \quad \int_{0}^{\infty} \phi(z)\,dz - \int_{0}^{z_2} \phi(z)\,dz = 0.08$$

Hence $$\int_{0}^{z_2} \phi(z)\,dz = \int_{0}^{\infty} \phi(z)\,dz - 0.08 = 0.5 - 0.08 = 0.42$$

From table III, $z_2 = 1.4$ ...(*iv*)

From (*i*) and (*ii*), $45 - \bar{x} = -0.5\sigma$

From (*iii*) and (*iv*), $64 - \bar{x} = 1.4\sigma$

Solving these equations, we get $\bar{x} = 50$ and $\sigma = 10$.

Example 26.48. *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*

(a) more than 2150 hours, *(b) less than 1950 hours and*

(c) more than 1920 hours and but less than 2160 hours. *(Bhopal, 2008 S ; U.P.T.U., 2008)*

Solution. Here $\mu = 2040$ hours and $\sigma = 60$ hours.

(*a*) For $x = 2150$, $z = \dfrac{x-\mu}{\sigma} = 1.833.$

∴ area against $z = 1.83$ in the table III = 0.4664.

We, however, require the area to the right of the ordinate at $z = 1.83$. This area = 0.5 – 0.4664 = 0.0336.

Thus the number of bulbs expected to burn for more than 2150 hours

$$= 0.0336 \times 2000 = 67 \text{ approximately.}$$

(*b*) For $x = 1950$, $z = \dfrac{x-\mu}{\sigma} = -1.5$

The area required in this case is to the left of $z = -1.33$

i.e., $= 0.5 - 0.4082$ (table value for $z = 1.33$)

$= 0.0918.$

∴ the number of bulbs expected to burn for less than 1950 hours

$$= 0.0918 \times 2000 = 184 \text{ approximately.}$$

(*c*) When $x = 1920$, $\quad z = \dfrac{1920 - 2040}{60} = -2$

When $x = 2160$, $\quad z = \dfrac{2160 - 2040}{60} = 2.$

The number of bulbs expected to burn for more than 1920 hours but less than 2160 hours will be represented by the area between $z = -2$ and $z = 2$. This is twice the area from the table for $z = 2$, *i.e.*, $= 2 \times 0.4772 = 0.9544$.

Thus the required number of bulbs = 0.9544 × 2000 = 1909 nearly.

Example 26.49. *If the probability of committing an error of magnitude x is given by*

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2};$$

compute the probable error from the following data :

$m_1 = 1.305$; $m_2 = 1.301$; $m_3 = 1.295$; $m_4 = 1.286$;
$m_5 = 1.318$; $m_6 = 1.321$; $m_7 = 1.283$; $m_8 = 1.289$;
$m_9 = 1.300$; $m_{10} = 1.286$. (*Kurukshetra, 2005*)

Solution. From the given data which is normally distributed, we have

$$\text{mean} = \frac{1}{10}\Sigma m_i = \frac{12.984}{10} = 1.2984$$

and

$$\sigma^2 = \frac{1}{10}\,\Sigma(m_i - \text{mean})^2$$

$$= \frac{1}{10}\,[(0.007)^2 + (0.003)^2 + (0.003)^2 + (0.012)^2 + (0.02)^2 + (0.023)^2 + (0.015)^2 + (0.009)^2 + (0.002)^2 + (0.012)^2]$$

$$= 0.0001594 \text{ whence } \sigma = 0.0126.$$

∴ *probable error* $= \dfrac{2}{3}\sigma = 0.0084$ approx.

Example 26.50. *Fit a normal curve to the following distribution.*

x :	*2*	*4*	*6*	*8*	*10*
f :	*1*	*4*	*6*	*4*	*1*

(*V.T.U., 2001*)

Solution.

$$\text{Mean} = \frac{\Sigma fx}{\Sigma f} = \frac{2 + 16 + 36 + 32 + 10}{16} = 6$$

$$\text{S.D.} = \sqrt{\left[\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2\right]} = \sqrt{(40 - 36)} = 2$$

Taking $\mu = 6$, $\sigma = 2$ and $N = 16$, the equation of the normal curve is

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \text{ or } y = \frac{1}{2\sqrt{2\pi}} e^{-(x-6)^2/8} \qquad ...(i)$$

Area under (*i*) in (x_1, x_2) or (z_1, z_2)

$$= \frac{1}{\sqrt{2\pi}} \int_0^{z_2} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2} dz \qquad \text{where } z = \frac{x-6}{2}$$

To evaluate these integrals, we refer to table III.

Calculations :

Mid x	(x_1, x_2)	(z_1, z_2)	*Area under (i) in* (z_1, z_2)	*Expected frequency*
2	(1, 3)	(– 2.5, – 1.5)	0.4938 – 0.4332	16 × 0.606 = 0.97
4	(3, 5)	(– 1.5, – 0.5)	0.4332 – 0.1915	16 × 0.2417 = 3.9
6	(5, 7)	(– 0.5, 0.5)	0.1915 + 0.1915	16 × 0.383 = 6.1
8	(7, 9)	(0.5, 1.5)	0.4332 – 0.1915	16 × 0.2417 = 3.9
10	(9, 11)	(1.5, 2.5)	0.4938 – 0.4332	16 × 0.606 = 0.97

Hence the expected (theoretical) frequencies corrected to nearest integer are 1, 4, 6, 4, 1 which agree with the observed frequencies. This shows that the normal curve (*i*) is a proper fit to the given distribution.

PROBLEMS 26.7

1. Show that the standard deviation for a normal distribution is approximately 25% more than the mean deviation.
2. For a normally distributed variate with mean 1 and S.D. 3, find the probabilities that
   (*i*) $3.43 \le x \le 6.19$ (*ii*) $-1.43 \le x \le 6.19$.
3. If z is normally distributed with mean 0 and variance 1, find
   (*i*) $P_r\,(z \le -1.64)$; (*ii*) z_1 if $P_r(z \ge z_1) = 0.84$.
4. In a certain examination, the percentage of candidates passing and getting distinctions were 45 and 9 respectively. Estimate the average marks obtained by the candidates, the minimum pass and distinction marks being 40 and 75 respectively. (Assume the distribution of marks to be normal). (*Kottayam, 2005*)
5. A manufacturer of air-mail envelopes knows from experience that the weight of the envelopes is normally distributed with mean 1.95 gm and standard deviation 0.05 gm. About how many envelopes weighing (*i*) 2 gm or more ; (*ii*) 2.05 gm or more can be expected in a given packet of 100 envelopes.
6. The mean height of 500 students is 151 cm. and the standard deviation is 15 cm. Assuming that the heights are normally distributed, find how many students' heights lie between 120 and 155 cm. (*Burdwan, 2003*)
7. The mean and standard deviation of the marks obtained by 1000 students in an examination are respectively 34.4 and 16.5. Assuming the normality of the distribution, find the approximate number of students expected to obtain marks between 30 and 60.
8. In an examination taken by 500 candidates, the average and the standard deviation of marks obtained (normally distributed) are 40% and 10%. Find approximately
   (*i*) how many will pass, if 50% is fixed as a minimum ?
   (*ii*) what should be the minimum if 350 candidates are to pass ?
   (*iii*) how many have scored marks above 60% ?
9. The mean inside diameter of a sample of 200 washers produced by a machine is 5.02 mm and the standard deviation is 0.05 mm. The purpose for which these washers are intended allows a maximum tolerance in the diameter of 4.96 to 5.08 mm, otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine, assuming the diameters are normally distributed.
   [**Hint.** 4.96 in standard units = (4.96 – 5.02)/0.05 = – 1.2
   5.08 in standard units = (5.08 – 5.02)/0.05 = 1.2
   Proportion of non-defective washers = 2 (area between $z = 0$ and $z = 1.2$)
   = 0.7698 or 77% nearly.
   ∴ percentage of defective washers = 100 – 77 = 23%.]
10. 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 0.0020 cm., how many of the plugs are likely to be rejected if the approved diameter is 0.752 ± 0.004 cm. ? (*Bhopal, 2002*)

11. It is given that the age of thermostats of a particular make follow the normal law with mean 5 years and S.D. 2 years. 1000 units are sold out every month. How many of them will have to be replaced at the end of the second year.

12. The income of a group of 10,000 persons was found to be normally distributed with mean Rs. 750 p.m., and standard deviation of Rs. 50. Show that, of this group, about 95% had income exceeding Rs. 668 and only 5% had income exceeding Rs. 832. Also find the lowest income among the richest 100. *(U.P.T.U., 2004 S)*

13. Find the equation of the best fitting normal curve to the following distribution :

x :	0	1	2	3	4	5
y :	13	23	34	15	11	4

14. Obtain the equation of the normal probability curve that may be fitted to the following data :

Variable :	4	6	8	10	12	14	16	18	20	22	24
Frequency:	1	7	15	22	35	43	38	20	13	5	1

15. A factory turns out an article by mass production and it is found that 10% of the product is rejected. Find the S.D. of the number of rejects and the equation to the normal curve to represent the number of rejects.

[**Hint.** $p = 0.1, q = 0.9, n = 100$.

$\therefore$ binomial distribution of rejects gives mean $= np = 10$, S.D. $= \sqrt{(npq)} = 3$

If this binomial distribution is approximated by a normal distribution, then the equation to the normal curve is

$$y = \frac{100}{\sigma\sqrt{(2\pi)}} e^{-(x-\mu)^2/2\sigma^2} \quad \text{where } \mu = 10, \sigma = 3.]$$

16. Given that the probability of committing an error of magnitude x is

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}, \text{ show that the probable error is } 0.4769/h.$$

26.18 NORMAL APPROXIMATION TO BINOMIAL DISTRIBUTION

If the number of successes in a Binomial distribution range from x_1 to x_2, then the probability of getting these successes

$$= \sum_{r=x_1}^{x_2} {}^nC_r\, p^r\, q^{n-r}$$

As the number of trials increases, the Binomial distribution becomes approximated to the Normal distribution. The mean np and the variance npq of the binomial distribution will be quite close to the mean and standard deviation of the approximated normal distribution. Thus *for n sufficiently large (≥ 30), the binomial distribution with probability of success p, is approximated by the normal distribution with* $\mu = np, \sigma = \sqrt{npq}$.

We must however, be careful to get the correct values of z. For any success x, real class interval is $(x - 1/2, x + 1/2)$. Hence

$$z_1 = \frac{x_1 - \frac{1}{2} - \mu}{\sigma} = \frac{x_1 - \frac{1}{2} - np}{\sqrt{npq}}\,;\; z_2 = \frac{x_2 + \frac{1}{2} - np}{\sqrt{npq}}$$

so that $P(x_1 < x < x_2) = P(z_1 < z < z_2) = \int_{z_1}^{z_2} \phi(z)\, dz$ which can be calculated by using table III–Appendix 2.

Example 26.51. *In a referendum 60% of voters voted in favour. A random sample of 200 voters was selected. What is the probability that in the sample*

(a) more than 130 voted in favour ?

(b) between 105 and 130 inclusive voted in favour ?

(c) 120 voted in favour ?

Solution. Here $n = 200, p = 0.6, q = 0.4$

$$\therefore \quad \mu = np = 200 \times 0.6 = 120;\ \sigma = \sqrt{npq} = \sqrt{48} = 6.928$$

$$(a)\ P(x > 130) = P(x > 130.5) = P\left(x > \frac{130.5 - 120}{\sqrt{48}}\right) = P(z > 1.516) = 0.0648$$

(b) $P(105 < x < 130) = P(105.5 < x < 129.5)$

$$= P\left(\frac{105.5 - 120}{\sqrt{48}} < z < \frac{129.5 - 120}{\sqrt{48}}\right) = P(-2.09 < z < 1.37) = 0.8964$$

(c) $P(x = 120) = P(119.5 < x < 120.5)$

$$= P(-0.072 < z < 0.072) = 0.0575.$$

PROBLEMS 26.8

1. A pair of unbiased dice are rolled 180 times and their score recorded. Find
(a) $P(x \le 20)$, (b) $P(20 \le x \le 40)$, (c) $P(20 < x \le 30)$.
2. A marksman has a probability of 0.9 of hitting a target on a single shot. If the marksman has 40 shots, what is the probability that he hits the target (a) at least 35 times; (b) between 34 and 36 times; (c) 37 times.
3. A certain drug is effective in 72% of cases. Given 2000 people are treated with the drug, what is the probability that it will be effective for (a) at least 1400 patients, (b) less than 1390 patients, (c) 1420 patients.

26.19 SOME OTHER DISTRIBUTIONS

Discrete distributions

(1) Geometric distribution. If p be the probability of success and k be the numbers of failures preceding the first success then this distribution is

$$P(k) = q^k p, \qquad k = 0, 1, 2, ..., q = 1 - p.$$

Obviously $$\sum_{k=0}^{\infty} P(k) = p \sum_{k=0}^{\infty} q^k = p \cdot \frac{1}{1-q} = 1.$$

It can easily be shown that mean = q/p, and variance = q/p^2.

(2) Negative binomial distribution. This distribution gives the probability that the event occurs for the kth time on the rth trial ($r \ge k$). If p be the probability of occurrence of an event then

$$P(k, r) = {}^{r-1}C_{k-1} p^k q^{r-k}.$$

It contains two parameters p and k. If $k = 1$, the Negative binomial distribution reduces to the geometric distribution.

(3) Hypergeometric distribution. Suppose a bag contains m white and n black balls. If r balls are drawn one at a time (*with replacement*), then the probability that k of them will be white is

$$P(k) = {}^mC_k\, {}^nC_{r-k} / {}^{m+n}C_r,\ k = 0, 1, ... r, r \le m, r \le n.$$

This distribution is known as *Hypergeometric distribution.*

For $$\sum_{k=0}^{r} P(k) = 1, \text{ since } \sum_{k=0}^{r} {}^mC_k\, {}^nC_{r-k} = {}^{m+n}C_r$$

This can be proved by equating the coefficient of t^r in

$$(1 + t)^m (t + 1)^n = (1 + t)^{m+n}$$

Continuous distributions

(4) Uniform (or Rectangular) distribution. A random variable X is said to be uniformly distributed over the interval $-\infty < a < b < \infty$, if its density is given by

$$f(x) = \frac{1}{b-a}, a < x < b \qquad ...(i)$$

The distribution given by (i) is called a *uniform distribution*. In this distribution, X takes the values with the same probability.

Its *mean* $$\mu = \int_a^b x \cdot f(x)\, dx = \frac{1}{b-a}\left|\frac{x^2}{2}\right|_a^b = \frac{a+b}{2}.$$

and *variance* $$\sigma^2 = \mu_2' - (\mu)^2 = \int_a^b x^2 \cdot \frac{1}{b-a}\, dx - \left(\frac{a+b}{2}\right)^2 = \frac{1}{12}(b-a)^2.$$

(5) Gamma distribution. This continuous distribution is given by $f(x) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x}$ for all $x \geq 0$, where r and λ (both > 0) are called the parameters of the *gamma distribution*. Its mean $= r/\lambda$ and variance $= r/\lambda^2$.

Gamma distribution tends to normal distribution as the parameter r tends to infinity.

(6) Exponential distribution. This distribution is a special case of gamma distribution when $r = 1$ so that $f(x) = \lambda e^{-\lambda x}$ for $x > 0$, where λ is a parameter.

It can be seen that mean $= 1/\lambda$, standard deviation $= 1/\lambda$.

This distribution plays an important role in the reliability and queuing theory.

(7) Weibull distribution*. This distribution is given by

$$f(x) = \frac{\alpha}{c} x^{\alpha-1} e^{-x^{\alpha}/c}, x > 0, c > 0$$

where c is a scale parameter and α a shape parameter.

Initially this distribution was used to describe experimentally observed variation in the fatigue resistance of steel and its elastic limits. But it has also been employed to study the variation of length of service of radio service equipment.

Example 26.52. *A die is cast until 6 appears. What is the probability that it must be cast more than 5 times?*

Solution. Here probability of getting 6 is $p = \frac{1}{6}$. Then $q = \frac{5}{6}$.

If X is the number of tosses required for the first success, then

$$P(X = x) = q^{x-1} p \text{ for } x = 1, 2, 3, \ldots$$

$\therefore$ required probability $= P(X > 5) = 1 - P(X \leq 5)$

$$= 1 - \sum_{x=1}^{5} \left(\frac{5}{6}\right)^{x-1} \cdot \left(\frac{1}{6}\right) = 1 - \frac{1}{6}\left\{1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^3 + \left(\frac{5}{6}\right)^4\right\} = \left(\frac{5}{6}\right)^5.$$

Example 26.53. *A random variable X has a uniform distribution over (– 3, 3), find k for which*

$$P(X > k) = \frac{1}{3}.$$

Also evaluate $P(X < 2)$ *and* $P[|X - 2| < 2]$.

Solution. (*i*) Density of $X = f(x) = \frac{1}{b-a} = \frac{1}{3-(-3)} = \frac{1}{6}$

$$\therefore \quad P(X > k) = 1 - P(X \leq k) = 1 - \int_{-3}^{k} f(x)\, dx$$

$$= 1 - \frac{1}{6}\int_{-3}^{k} dx = 1 - \frac{1}{6}(k+3) = \frac{1}{3} \qquad \text{(given)}$$

This gives $k = 1$.

(*ii*) $$P(X < 2) = \int_{-3}^{2} f(x)\, dx = \frac{1}{6}\int_{-3}^{2} dx = \frac{5}{6}.$$

(*iii*) $$P[|X-2| < 2] = P[2-2 < X < 2+2] = P[0 < x < 4] = \int_{0}^{3} f(x)\, dx = \frac{1}{6}\int_{0}^{3} dx = \frac{1}{2}.$$

PROBLEMS 26.9

1. Show that the mode of the *geometric distribution* $P(x) = \left(\frac{1}{2}\right)^x$, $x = 1, 2, 3 \ldots$, is unity.
2. Show that for the *rectangular distribution* $f(x) = 1, 0 \leq x \leq 1$,

$$\text{mean} = \frac{1}{2}, \text{ variance} = \frac{1}{12} \text{ and mean deviation} = \frac{1}{4}.$$

* It was first used by Swedish scientist *Weibull* in 1951.

3. Find the mean and variance of the *uniform distribution* given by $f(x) = 1/n, x = 1, 2, ..., n$.
4. Show that for the *exponential distribution*
$dP = y_0 e^{-x/\sigma}, 0 \leq x \leq \infty,$
the mean and S.D. are both equal to σ.
5. Find the mean and variance of the *exponential distribution* $f(x) = \frac{1}{b} e^{-(x-a)/b} x > a.$ *(Mumbai, 2005)*
6. Find the moment generating function for the *triangular distribution* given by
$$f(x) = x, \quad 0 \leq x \leq 1$$
$$= 2 - x, \quad 1 \leq x \leq 2.$$
7. Show that for the *Gamma distribution* $f(x) = \dfrac{e^{-x}x^{l-1}}{\Gamma(l)}$, $0 < x < \infty$, the mean and variance are both equal to l.
8. Find the moment generating function of the *Gamma distribution* $f(x) = \dfrac{1}{\Gamma\left(\frac{1}{4}\right)} e^{-x} x^{-3/4}, x \geq 0$, at the origin.

(J.N.T.U., 2006 ; Madras, 2000 S)

[**Chebyshev's inequality***. If x is a continuous random variable with mean μ and variance σ^2, then for any positive real parameter t,
$$P(|x-2| \geq t) \leq \sigma^2/t^2 \text{ or } P(|x-\mu| \leq t) \geq 1 - \sigma^2/t^2.$$
This result is known as *Chebyshev's inequality*. It gives limits to the probability that the value of the variate chosen at random will differ from mean by more than t.]

9. For the points on a symmetrical die, prove that *Chebyshev's inequality* gives
$$P(|x - \bar{x}| > 2.5) < 0.478,$$
while the actual probability is zero.
10. For the *Geometrical distribution* $P(x) = 2^{-x}, x = 1, 2, 3, ...$, prove that *Chebyshev's inequality* gives
$$P(|x-2| < 2) > \frac{1}{2},$$
while the actual probability is 15/16.

26.20 OBJECTIVE TYPE OF QUESTIONS

PROBLEMS 26.10

Select the correct answer or fill up the blanks in each of the following problems :

1. The probability that A happens is 1/3. The odds against happening of A are
(a) 2 : 1 (b) 2 : 3 (c) 3 : 2 (d) 5 : 2.
2. The odds in favour of an event A are 5 to 4. The probability of success of A is
(a) 4/5 (b) 5/9 (c) 4/9.
3. The probability that A passes a test is 2/3 and the probability that B passes the same test is 3/5. The probability that only one of them passes is
(a) 2/5 (b) 4/15 (c) 2/15 (d) 7/15.
4. A buys a lottery ticket in which the chance of winning is 1/10 ; B has a ticket in which his chance of winning is 1/20. The chance that atleast one of them wins is
(a) 1/200 (b) 29/200 (c) 30/200 (d) 170/200.
5. The probability that a non-leap year should have 53 Tuesdays is ...
6. The probability of getting 2 or 3 or 4 from a throw of single dice is ...
7. The mean of the Binomial distribution with n observations and probability of success p, is
(a) pq (b) np (c) $\sqrt{np}$ (d) $\sqrt{pq}$.
8. If the mean of a Poisson distribution is m, then S.D. of this distribution is
(a) m^2 (b) $\sqrt{m}$ (c) m (d) none of these.

* See footnote on page 571.

9. The S.D. of the Binomial distribution is
   (a) $\sqrt{npq}$ (b) $\sqrt{np}$ (c) npq (d) pq.
10. In a Poisson distribution if $2P(x = 1) = P(x = 2)$, then the variance is
   (a) 0 (b) – 1 (c) 4 (d) 2.
11. If the probability of hitting a target by one shot be $p = 0.8$, then the probability that out of ten shots, seven will hit the target is ...
12. For a Poisson variate x : $P(x = 1) = P(x = 2)$, then the mean of x is ...
13. If $P(A) = 0.35$, $P(B) = 0.73$ and $P(A \cap B) = 0.14$, then $P(A \cap B') = ...$
14. If A and B are independent, $P(B) = 0.14$ and $P(A/B) = 0.24$, then $P(A) = ...$
15. The probability distribution of the number of heads, when two coins are tossed, is ...
16. The multiplication law of probability states that ...
17. The area under the standard normal curve which lies between $z = 0.90$ and $z = -1.85$ is ...
   [Given $P(0 < z < 1.85) = 0.4678$, $P(0 < z < 0.9) = 0.3159$]
18. The mean, median and mode of a normal distribution are ...
19. The mean and variance of a Poisson distribution are ...
20. If A and B are two mutually exclusive events, then $P(A \cup B) = ...$
21. For a normal distribution $\beta_1 = ...$ and $\beta_2 = ...$
22. The number of ways in which five people can be lined up to get on a bus are ...
23. A shipment of 10 television sets contains 3 defective sets. The number of ways in which one can purchase 4 of these sets and receive 2 defective sets are ...
24. The probability of getting a total of 5 when a pair of dice is tossed is ...
25. If $P(B) = 0.81$ and $P(A \cap B) = 0.18$, then $P(A/B) = ...$
26. If two unbiased dice are thrown simultaneously, the probability that the sum of the numbers on them is at least 10, is ....
27. If X is a Poisson variate such that $P(X = 2) = P(X = 3)$, then $P(X = 0) = ...$
28. An unbiased die is tossed twice, then the probability of obtaining the sum 6, is ...
29. The variance of Poisson distribution with parameter $\lambda = 2$ is ...
30. The distribution in which mean, median, mode are equal is ...
31. For the Poisson variate, probability of getting at least one success is ...
32. Total number of events in rolling of an ideal die is ...
33. If X be normal with mean 10 and variance 4, then $P(X < 11) = ...$
34. If X is a binomial variate with parameters n and p, then its m.g.f. about the origin is ...
35. In a normal distribution, mean deviation : standard deviation = ...
36. If A and B are independent and $P(A) = 1/2$, $P(B) = 1/3$ then $P(A \cap B) = ...$
37. If X is the random variable representing the outcome of the roll of an ideal die, then $E(X) = ...$
38. If X is a binomial variate with $p = 1/5$ for the experiment of 50 trials, then the standard deviation is ...
39. The area under the whole normal curve is ...
40. Given $X = B(n, p)$, then the conditions under which X tends to a Poisson distribution, are ...
41. If A and B are mutually exclusive events then $P(A \cup B) = ...$
42. The probability of selecting x white balls from a bag containing y white and z red balls is ...
43. The mean of the binomial distribution is ...
44. If A and B are mutually exclusive events, $P(A) = 0.29$, $P(B) = 0.43$, then $P(A \cup B) = ...$ and $P(A \cap B') = ...$
45. If the mean and variance of a binomial variate are 12 and 4, then the distribution is ...
46. If x is a Poisson variable such that $P(x = 2) = 9\,P(x = 4) + 90\,P(x = 6)$, then the mean = ...
47. μ_r' the rth moment about the origin in terms of the m.g.f. is ...
48. The chance of throwing 7 in a single throw with two dice is ...
49. If A and B are any two events with $P(A) = 1/2$, $P(B) = 1/3$ and $P(A \cap B) = 1/4$, then $P(A/B) = ...$
50. In the roll of an ideal die, the probability of getting a prime number is ...
51. If A and B are mutually exclusive events, $P(A \cup B) = 0.6$, $P(B) = 0.4$, then $P(A) = ...$
52. The probability that a leap year should have 53 Sundays is ...
53. The probability density function of a binomial distribution is ...
54. The probable error is ... times S.D. approximately.
55. To fit a normal distribution, the parameters required are ...

56. A card is drawn from a well-shuffled pack of 52 cards, then the probability of this card being a red coloured ace is ...
57. If $P(1) = P(2)$, then the mean of the Poisson distribution is ...
58. Baye's theorem states that ...
59. If x is a Poisson variate such that $P(x = 1) = 0.3$ and $P(x = 2) = 0.2$, then $P(x = 0) = ...$
60. If A and B are events such that $P(A \cup B) = 3/4$, $P(A \cap B) = 1/4$, and $P(A) = 2/3$, then $P(A) = ...$
61. The chance of throwing 7 in a single throw with two dice is the same as that of getting 7 in two throws of a single die. (True or False)
62. If the mean of a Poisson distribution is 5, then its variance is 10. (True or False)
63. If X is normal with mean 3 and variance 1, then $X - 3$ is a variate with mean 0 and variance 1. (True or False)
64. If X is a binomial variable with parameters $n = 10$, $p = 1/4$, then its standard deviation is 2.25. (True or False)
65. The mean of a binomial distribution is 5 and S.D. is 3. (True or False)
66. The mean and variance of Poisson distribution are equal. (True or False)
67. The graph of the normal distribution is symmetric with respect to the line $y = x$. (True or False)
68. The standard deviation of a binomial distribution is np. (True or False)
69. $f(x) = kx$ in $0 < x < 1$ is a valid probability density function, if $k = ...$
70. If $V(x) = 2$, then $V(2x + 3) = ...$
71. The p.d.f. of an exponential distribution is ...
72. If X is uniformly distributed in $(-2, 3)$, then its variance is ...
73. The variance of Poisson distribution with parameter $\lambda = 2$ is ...
74. In Gamma distribution with parameter l, the variance is ...
75. If $f(x) = kx^3$, $0 < x < 1$ and 0 elsewhere, is a p.d.f., then $k = ...$
76. The m.g.f. of a random variable X is $(1 - 2t)^{-4}$, then $E(x)$ is ...
77. A random variable X has F-distribution with (m, n) degrees of freedom, then $1/X$ has the same distribution with ... degrees of freedom.
78. If X is a continuous random variable having the p.d.f. $f(x)$, then the m.g.f. about the origin is given by ...
79. If $f(X) = X + 2/k$, $X = 1, 2, 3, 4, 5$ is the probability distribution of a discrete random variable, then $k = ...$
80. The p.d.f. of Gamma variate is ...
81. If X is uniformly distributed in $[a, b]$, then $E(X) = ...$
82. The marks obtained by students were found normally distributed with mean 75 and variance 100. The percentage of students who scored more than 75 marks is ...
83. When four unbiased coins are tossed, the probability of getting two heads is ...
84. If $f(x) = \begin{cases} kxe^{-x}, & x > 0 \\ 0, & \text{elsewhere,} \end{cases}$ is the p.d.f. of x, then $k = ...$
85. If x is a uniform distribution defined in the interval (4, 7), then its variance is ...
86. The p.d.f. of a continuous random variable is $f(x) = k/x^3$, $5 \le x \le 10$; 0 elsewhere, then the value of k is
    (a) 1 (b) 50 (c) 200/3 (d) 200.
87. The relation between probability density function and cumulative density function of a random variable is ...
88. If X has Poisson distribution with parameter λ, then P (X is even) = ...
89. Range of t- distribution is ...
90. If the p.d.f. of x is $f(x) = kx(1 - x)$, $0 < x < 1$, then $k = ...$
91. The function $f(x) = \begin{cases} kx^2, & 0 < x < 3 \\ 0, & \text{otherwise,} \end{cases}$ is a probability function, then $k = ...$
92. If the random variable x is uniformly distributed in [0, 3], then its p.d.f. is $f(x) = 3$, $0 < x < 3$; 0, elsewhere. (True or False)
93. Exponential distribution $f(x)$ is defined by $f(x) = ae^{-2x}$, $0 < x < \infty$, then $a = ...$
94. The p.d.f. of Beta distribution with $\alpha = 1$, $\beta = 4$ is $f(x) = ...$
95. For a standard normal variate z, $P(-0.72 \le z \le 0) = ...$