Chapter 6

. . . is touched by that dark miracle of chance which makes new magic in a dusty world.

—Thomas Wolfe

Fundamentals of Probability

LEARNING OBJECTIVES

After studying this chapter, you should be able to

- understand the amount of uncertainty that is involved before making important decisions.
- understand fundamentals of probability and various probability rules that help you to measure uncertainty involving uncertainty.
- perform several analyses with respect to business decision involving uncertainty.

6.1 INTRODUCTION

Statistical methods, discussed so far, are helpful in summarizing sample data to gain knowledge about the entire population or process. However, decision makers often come across some degree of risk while selecting a particular course of action or strategy to solve a decision problem involving uncertainty. It is because each strategy can lead to a number of different possible outcomes (or results). Thus, it is necessary for the decision makers to gain knowledge of the concepts of probability, and probability distributions to choose a course of action to arrive at an optimal decision.

Random Experiment: A process of obtaining information through observation or measurement of a phenomenon whose outcome is subject to chance.

6.2 CONCEPTS OF PROBABILITY

The definition of following terms and their importance in probability theory will be useful for decision makers to obtain a deeper understanding of the subject.

6.2.1 Random Experiment

Random experiment (also called *act, trial, operation* or *process*) is an activity that leads to the occurrence of one and only one of several possible outcomes that is not predictable until the completion of the experiment. The *random experiment* has following properties: (i) all possible outcomes can be listed in advance, (ii) experiment can be repeated, and (iii) the same outcome may not occur on various repetitions. The variation among experimental outcomes due to uncontrolled factors is called ***random variation***. It is assumed that the effects of uncontrolled factors vary randomly and unpredictably during repetition of the experiment. Each random experiment may result in one or more outcomes, also called **events** and denoted by capital letters

The outcome (event or observation) of an experiment may be expressed in numerical or non-numerical value. Few examples of both types are as follows:

Numerical outcome

- Counting the number of arrivals at a service window.
- Measuring blood pressure of a group of individuals.
- Checking an automobile's petrol consumption per km.

Non-numerical outcome

- Tossing a coin and observing the face, head or tail.
- Testing quality of a product to know defective or acceptable.
- Payment made by cash, cheque or credit card.

In all such cases, the outcome will not be known with certainty in advance until experiment is not over. Also the number of outcomes may be finite or infinite depending on the nature of the experiment. For example, in the experiment of tossing a coin, the outcomes are finite and represented by the head and tail, whereas in the experiment of measuring the time between successive breakdowns of a machine, the outcomes are infinite and represented by the time of breakdown. Although an individual outcome associated with a random experiment cannot be predicted exactly, the frequency of occurrence of such an outcome can be recorded in a large number of repetitions and becomes the basis for solving problems dealing with uncertainty.

A Simple Event: The basic possible outcome of an experiment, it cannot be broken down into simple outcomes.

6.2.2 Sample Space

The set of all distinct outcomes (events) for a random experiment is called the **sample space** (or *event space*) provided.

- Two or more of these outcomes do not occur simultaneously.
- Each random experiment is resulting into exactly one of the outcomes.

Sample space is denoted by the capital letter S.

Sample Space: The set of all possible outcomes or simple events of an experiment.

Illustrations

1. Consider the experiment of recording category of blood. The following four possible outcomes represent simple events:

   E_1 : Blood type A E_2 : Blood type B

   E_3 : Blood type AB E_4 : Blood type O

   The sample space is $S = \{E_1, E_2, ..., E_8\}$. The sample events can be displayed in a *tree diagram* as shown in Fig. 6.1.

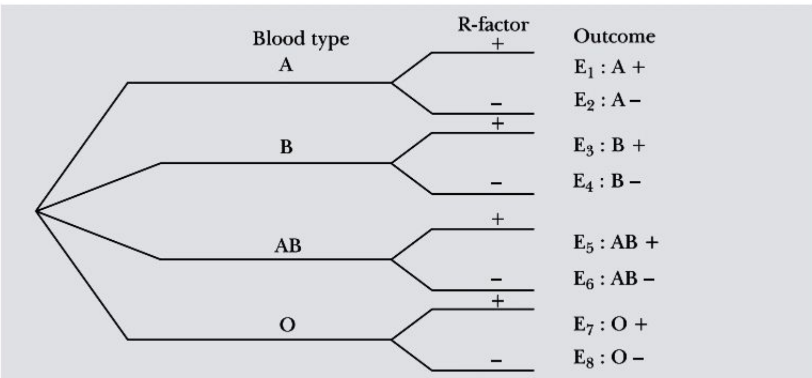


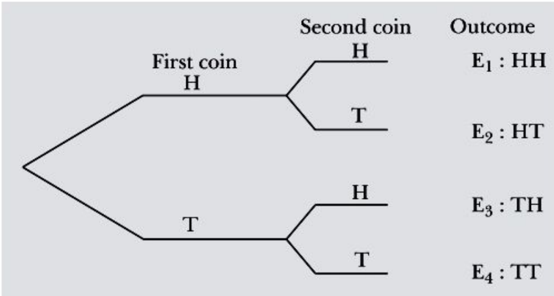
Figure 6.1
Tree Diagram

2. Consider the experiment of tossing two coins. The four possible outcomes are the following simple events.

$$E_1 : HH \quad E_2 = HT \quad E_3 : TH \quad E_4 : TT$$

The sample space is S = {E_1, E_2, E_3, E_4}. The sample events can be displayed in a tree diagram as shown in Fig. 6.2.

Figure 6.2
Tree Diagram



6.2.3 Types of Event

Event: Any subset of outcomes of an experiment.

A single possible outcome (or result) of an experiment is called a ***simple* (*or elementary*) *event***. An ***event*** is the set (or collection) of one or more simple events of an experiment in the sample space and having a specific common characteristic. For example, for the above defined sample space S, the collection (H, T), (T, H) is the event containing simple event as H or T. Other examples of events are as follows:

- More than 5 customers at a service facility in one hour.
- 75 per cent marks or better in an examination.
- Sales volume of a retail store more than ₹2,00,000 on a given day.

Mutually Exclusive Events

Mutually Exclusive Events: Events which cannot occur together or simultaneously.

If two or more events cannot occur simultaneously in a single trial of an experiment, then such events are called mutually exclusive events or disjoint events. In other words, two events are mutually exclusive if the occurrence of one of them prevents or rules out the occurrence of the other. For example, the numbers 2 and 3 cannot occur simultaneously on the roll of a dice.

Symbolically, a set of events {$A_1, A_2, ..., A_n$} is mutually exclusive if $A_i \cap A_j = \varnothing \, (i \neq j)$. This means the intersection of two events is a null set (∅).

Collectively Exhaustive Events

Collectively Exhaustive Events: The list of events that represents all possible experimental outcomes.

A list of all possible events that can occur from an experiment is said to be collectively exhaustive events. Symbolically, a set of events {$A_1, A_2, ..., A_n$} is collectively exhaustive when union of these events is identical with the sample space S. That is,

$$S = \{A_1 \cup A_2 \cup \ldots \cup A_n\}$$

For example, the number 7 cannot come upon the uppermost face during the experiment of rolling a dice because the number uppermost faces form the sample space S = {1, 2, 3, 4, 5, 6}.

Independent and Dependent Events

Two or more events are said to be *independent* when one event does not affect, and is not affected by other event. For example, (i) outcomes of successive tosses of a coin are independent of outcomes in the preceding toss and (ii) increase in the population (in per cent) per year in India is independent of increase in wheat production (in per cent) per year in the USA.

Two or more events are said to be *dependent* when occurrence of one event affects or get affected by other event. For example, drawing of a card (say a queen) from a pack of playing cards without replacement reduces the chances of drawing a queen in the subsequent draws.

Compound Events

Two or more events are said to be compound events when these events occur simultaneous. These events may be (i) independent or (ii) dependent.

Equally Likely Events

Two or more events are said to be equally likely when each of these has an equal chance to occur. That is, one of them cannot be expected to occur in preference to the other. For example, each number may be expected to occur on the uppermost face of a rolling die the same number of times in the long run.

Complementary Events

If E is any subset of the sample space, then its complement denoted by E (read as E-bar) contains all the elements of the sample space that are not part of E. If S denotes the sample space, then

$$\begin{aligned}\overline{E} &= S - E \\ &= \{\text{All sample elements not in E}\}\end{aligned}$$

For example, if E represents companies with sales less than or equal to ₹25 lakh written as $E = \{x : x \leq 25\}$, then this set is a complement of the set, $\overline{E} = \{x : x > 25\}$. Obviously such events must be mutually exclusive and collective exhaustive.

6.3 DEFINITION OF PROBABILITY

Probability: A numerical measure of the likelihood of occurrence of an uncertain event.

A general definition states that the **probability** is a numerical measure (between 0 and 1 inclusively) of the likelihood or chance of occurrence of an uncertain event. Following three approaches of calculating probability of an event are as follows:

6.3.1 Classical Approach

Classical Approach: The probability of an event A is the ratio of the number of outcomes in favour of A to the number of all possible outcomes, provided experimental outcomes are equally likely to occur.

This approach of defining probability is based on the assumption that all possible outcomes (finite in number) of an experiment are mutually exclusive and equally likely.

If a random experiment is repeated finite number of times, out of which outcomes '*a*' are in favour of event A, outcomes '*b*' are not in favour of event A and all these possible outcomes are mutually exclusive, collectively exhaustive and equally likely, then probability of occurrence of event A is defined as:

$$P(A) = \frac{a}{a+b} = \frac{\text{Number of favourable outcomes}}{\text{Total number of possible outcomes}} = \frac{c(A)}{c(S)}$$

For example, if a die is rolled, then on any trial, each outcome or event (face or number) is equally likely to occur. Since there are six equally likely exhaustive outcomes, therefore probability of any one event (face or number) occurring is 1/6. In general, for a random experiment with n mutually exclusive, exhaustive, equiprobable events, the probability of any of the events is equal to $1/n$.

Since the probability of occurrence of an event is based on prior knowledge of the process involved, therefore this approach is often called *a priori (original) approach or classical approach*. This approach implies that there is no need to perform random experiments to find the probability of occurrence of an event. Also, no experimental data are required for computation of probability.

Since assumption that the outcomes are equally likely may not be verified with certainty, therefore this approach of calculating probability is not used often other than in games of chance. For example, if two children in a family are classified according to their sex, then the possible outcomes for number of boys are 0, 1, 2. Thus, according to the classical approach, the probability for each of the outcomes should be one-third. But the probability for each of the outcome is 1/4, 1/2 and 1/4 for 0, 1 and 2 boys, respectively.

6.3.2 Relative Frequency Approach

Relative Frequency Approach: The probability of an event A is the ratio of the number of times that A has occurred in n trails of an experiment.

If outcomes or events of a random experiment are not equally likely or not known whether they are equally likely, then classical approach is not desirable to determine probability of a random event. For example, in cases like (i) whether a number greater than 3 will appear when die is rolled or (ii) whether a lot of 100 items will contain 10 defective items, etc., it is not possible to predict occurrence of an outcome in advance without repetitive trials of the experiment.

This approach of computing probability states that when a random experiment is repeated a large number of times under identical conditions where trials are independent to each other, the desired event may occur some proportion (relative frequency) of time. Thus, probability of an event can be approximated by recording the relative frequency with which such an event has occurred over a finite number of repetitions of the experiment under identical conditions.

For example, if a die is tossed n times and s denotes the number of times the event A (say, number 4, 5, or 6) occurs, then the ratio $P(A) = c(s)/n$ gives the proportion of times the event A occurs in n trials, and are also called relative frequencies of the event in n trials. Although the estimate about P(A) may change after every trial, yet the proportion $c(s)/n$ tends to cluster around a unique central value as the number of trials n becomes even larger. This unique central value (also called probability of event A) is defined as

$$P(A) = \lim_{n \to \infty} \left\{ \frac{c(s)}{n} \right\}$$

where $c(s)$ represents the number of times that an event s occurs in n trials of an experiment.

Since the probability of an event is determined through repetitive empirical observations of experimental outcomes, it is also known as *empirical probability*. Few situations to which this approach can be applied are follows:

- Observing how often you win lottery when buying regularly.
- Observing whether or not a certain traffic signal is red when you cross it.
- Observing births and noting how often the baby is a female.

6.3.3 Subjective Approach

Subjective Approach: The probability of an event based on the personal beliefs of an individual.

The **subjective approach** of calculating probability is always based on the degree of beliefs, convictions and experience concerning the likelihood of occurrence of a random event. It is a way to quantify an individual's beliefs, assessment and judgment about a random phenomenon.

Probability assigned for the occurrence of an event may be based on just guess or on having some idea about the relative frequency of past occurrences of the event. This approach must be used when either sufficient data are not available or sources of information giving different results are not known.

6.3.4 Fundamental Rules of Probability

No matter which approach is used to define probability, the following fundamental rules must be satisfied. Let S be the sample space of an experiment containing, n mutually exclusive and exhaustive events (either elementary or compound) $A_1, A_2, \ldots, A_n$. The **fundamental rules of** probability of any event A in S are as follows:

- Each probability should fall between 0 and 1, i.e. $0 \leq P(A_i) \leq 1$, for all i, i.e., probability of an event is restricted to the range *zero* to *one(both included)*, where zero represents an impossible event and one represents a certain event.
- $P(S) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1$, i.e., the sum of probabilities of all simple events in a sample space is equal to one.

- Probability of an impossible event or an empty set is zero, i.e. $P(\Phi) = 0$.
- If events A_1 and A_2 are two elements in S and if occurrence of A_1 implies that of A_2 occurs (A_1 is a subset of A_2), then the probability of A_1 is less than or equal to the probability of A_2. That is, $P(A_1) \leq P(A_2)$.
- Probability of an event that does not occur is equal to one minus the probability of the event that does occur (rule for complementary events), i.e. $P(\overline{A}) = 1 - P(A)$.

6.3.5 Glossary of Probability Terms

If A and B are two events, then

$A \cup B$ = an event which represents the occurrence of either A or B or both.

$A \cap B$ = an event which represents the simultaneous occurrence of A and B.

$\overline{A}$ = complement of event A and represents non-occurrence of A.

$\overline{A} \cap \overline{B}$ = both A and B do not occur.

$\overline{A} \cap B$ = event A does not occur but event B occurs.

$A \cap \overline{B}$ = event A occurs but event B does not occur.

$(A \cap \overline{B}) \cup (\overline{A} \cap B)$ = exactly one of the two events A and B occurs.

6.4 COUNTING RULES FOR DETERMINING THE NUMBER OF OUTCOMES

In order to assign probabilities to experimental outcomes, it is first necessary to identify and count them. The following are three important rules for counting the experimental outcomes.

6.4.1 Multistep Experiments

The counting rule for multistep experiments helps us to determine the number of experimental outcomes without listing them. The rule is defined as

If an experiment is performed in k stages with n_1 ways to accomplish the first stage, n_2 ways to accomplish the second stage... and n_k ways to accomplish the kth stage, then the number of ways to accomplish the experiment is $n_1 \times n_2 \times \ldots \times n_k$.

Illustrations

1. Tossing of two coins can be thought of as a two-step experiment in which each coin can land in one of two ways: head (H) and tail (T). Since the experiment involves two steps, forming the pair of faces (H or T), the total number of simple events in S will be

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

The elements of S indicate that there are $2 \times 2 = 4$ possible outcomes.

When the number of alternative events in each of the several trials is same, that is, $n_1 = n_2 = \ldots = n_k$, then the multistep method gives $n_1 \times n_2 \times \ldots \times n_k = n^k$.

For example, if the coins involved in a coin-tossing experiment are four, then the number of experimental outcomes will be $2 \times 2 \times 2 \times 2 = 24 = 16$.

2. Suppose a person can take three routes from city A to city B, four from city B to city C and three from city C to city D. Then the possible routes for reaching from city A to D, while he must travel from A to B to C to D are (A to B) × (B to C) × (C to D) $= 3 \times 4 \times 3 = 36$ ways.

6.4.2 Combinations

Sometimes, objects that are chosen are important rather than the ordering or arrangement of these. For example, (i) which books you are able to shelve is important rather than order of books in the shelve and (ii) selection of few students for a committee from a group of students is important rather than order of choice.

The counting rule for combinations allows to select r (say) number of outcomes from a collection of n distinct outcomes without taking into consideration the order of their arrangement. This rule is denoted by

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

where $n! = n(n-1)\,(n-2)\,...\,3.\,2.\,1$ and $0! = 1$.

The notation "!" means *factorial*, for example, $4! = 4 \times 3 \times 2 \times 1 = 24$.

Important Results

- ${}^nC_r = {}^nC_{n-r}$ and ${}^nC_n = 1$.
- If n objects consist of all n_1 of one type, all n_2 of another type, and so on up to n_k of the kth type, then the total number of selections that can be made of 1, 2, 3 up to n objects is $(n_1 + 1)\,(n_2 + 1)\,...\,(n_k + 1) - 1$.
- The total number of selections from n objects all different is $2^n - 1$.

6.4.3 Permutations

This rule of counting involves ordering or permutations and helps to compute the number of ways in which n distinct objects can be arranged, taking r of them at a time.

The total number of permutations of n objects taken r at a time is given by

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

Permuting each combination of r objects among themselves, we shall obtain all possible permutations of n objects, r at a time. Each combination gives rise to $r!$ permutations, so that $r!\; C(n, r) = P(n, r) = n!/(n-r)!$.

Example 6.1: Of 10 electric bulbs, three are defective but it is not known which are defective. In how many ways can three bulbs be selected? How many of these selections will include at least one defective bulb?

Solution: Three bulbs out of 10 bulbs can be selected in ${}^{10}C_3 = 120$ ways. The number of selections which include exactly one defective bulb will be ${}^7C_2 \times {}^3C_1 = 63$.

Similarly, the number of selections which include exactly two and three defective bulbs will be ${}^7C_1 \times {}^3C_2 = 21$ and ${}^3C_3 = 1$, respectively. Thus, the total number of selections including at least one defective bulb is $63 + 21 + 1 = 85$.

Example 6.2: A bag contains 6 red and 8 green balls.

(a) If one ball is drawn at random, then what is the probability of the ball being green?

(b) If two balls are drawn at random, then what is the probability that one is red and the other green?

Solution: (a) Since the bag contains 6 red and 8 green balls, therefore it contains 6 + 8 = 14 equally likely outcomes, that is, S = $\{r, g\}$. But one ball out of 14 balls can be drawn in ways, that is,

$${}^{14}C_1 = \frac{14!}{1!\,(14-1)!} = 14 \text{ ways}$$

Let A be the event of drawing a green ball. Then, out of these 8 green balls, one green ball can be drawn in ${}^8C_1$ ways:

$${}^8C_1 = \frac{8!}{1!\,(8-1)!} = 8$$

Hence,

$$P(A) = \frac{c(A)}{c(S)} = \frac{8}{14}$$

(b) All exhaustive number of cases, $c(S) = {}^{14}C_2 = \frac{14!}{2!\,(14-2)!} = 91$.

Also, out of 6 red balls, one red ball can be drawn in ${}^6C_1$ ways and out of 8 green balls, one green ball can be drawn in ${}^8C_1$ ways. Thus, the total number of favourable cases are:

$$c(B) = {}^6C_1 \times {}^8C_1 = 6 \times 8 = 48$$

Thus, $$P(B) = \frac{c(B)}{c(S)} = \frac{48}{91}$$

Example 6.3: Tickets are numbered from 1 to 100. They are well shuffled and a ticket is drawn at random. What is the probability that the drawn ticket has

(a) an even number?
(*b*) the number 5 or a multiple of 5?
(c) a number which is greater than 75?
(*d*) a number which is a square?

Solution: Since any of the 100 tickets can be drawn, therefore exhaustive number of cases are $c(S) = 100$.

(a) Let A be the event of getting an even numbered tickets. Then, $c(A) = 50$, and hence

$$P(A) = 50/100 = 1/2$$

(b) Let B be the event of getting a ticket bearing the number 5 or a multiple of 5, i.e.,

$$B = [5, 10, 15, 20,..., 95, 100]$$

which are 20 in number, $c(B) = 20$. Thus, $P(B) = 20/100 = 1/5$.

(c) Let C be the event of getting a ticket bearing a number greater than 75, i.e.,

$$C = \{76, 77,..., 100\}$$

which are 25 in number, $c(C) = 25$. Thus, $P(C) = 25/100 = 1/4$.

(*d*) Let D be the event of getting a ticket bearing a number which is a square, that is,

$$D = \{1, 4, 9, 16, 25, 36, 49, 64, 81, 100\}$$

which are 10 in number, $c(D) = 10$. Thus, $P(D) = 10/100 = 1/10$.

Conceptual Questions 6A

1. (a) Discuss the different schools of thought on the interpretation of probability. How does each school define probability?
   (b) Describe briefly the various schools of thought on probability. Discuss its importance in business decision-making.
   [*HP Univ., MBA, 2004; Delhi Univ., MBA, 2005*]
   (c) Examine critically the different schools of thought on probability.
   [*Delhi Univ., MBA, 2005; Kumaon Univ., 2006*]
2. Explain what you understand by the term probability. Discuss its importance in business decision making.
   [*Delhi Univ., MBA, 2002*]
3. (a) Give the classical and statistical definitions of probability and state the relationship, if any, between the two definitions.
   (b) Critically examine the '*a priori*' definition of probability showing clearly the improvement which the empirical version of probability makes over it.
4. Define independent and mutually exclusive events. Can two events be mutually exclusive and independent simultaneously? Support your answer with an example.
   [*Sukhadia Univ., MBA; Delhi Univ., MBA, 2005*]
5. Compare and contrast the three interpretations of probability.
6. Explain the difference between statistically independent and statistically dependent events.
7. Explain the meaning of each of the following terms:
   (a) Random phenomenon
   (b) Statistical experiment
   (c) Random event
   (d) Sample space
8. What do you mean by probability? Explain the importance of probability.
   [*Madras Univ., MA(Eco), MBA, 2003*]
9. State the multiplicative theorem of probability. How is the result modified when the events are independent.
10. Life insurance premiums are higher for older people, but auto insurance premiums are generally higher for younger people. What does this suggest about the risks and probabilities associated with these two areas of insurance business?
11. Distinguish between the two concepts in each of the following pairs:
    (a) Elementary event and compound events
    (b) Mutually exclusive events and overlapping events
    (c) Sample space and sample point
12. (a) Define the terms—joint probability, marginal probability and conditional probability

(b) By comparing the three kinds of probabilities (joint, conditional and marginal), explain what information is provided by each.

13. Suppose an entire shipment of 1000 items is inspected and 50 items are found to be defective. Assume the defective items are not removed from the shipment before being sent to a retail outlet for sale. If you purchase one item from this shipment, what is the probability that it will be one of the defective items?

14. Suppose you are told that the price of a particular stock will increase with a probability of 0.7.

(a) How is this probability interpreted?

(b) Assuming the definition of probability in terms of long-run relative frequencies, how would you find the probability that a stock price will increase?

Self-practice Problems 6A

6.1 Three unbiased coins are tossed. What is the probability of obtaining:

(a) all heads (b) two heads
(c) one head (d) at least one head
(e) at least two heads, (f) all tails

6.2 A card is drawn from a well-shuffled deck of 52 cards. Find the probability of drawing a card which is neither a heart nor a king.

6.3 In a single throw of two dice, find the probability of getting (a) a total of 11, (b) a total of 8 or 11, and (c) same number on both the dice.

6.4 Five men in a company of 20 are graduates. If 3 men are picked out of the 20 at random, what is the probability that they are all graduates? What is the probability of at least one graduate?

6.5 A bag contains 25 balls numbered 1 through 25. Suppose an odd number is considered a 'success'. Two balls are drawn from the bag with replacement. Find the probability of getting

(a) two successes (b) exactly one success
(c) at least one success (d) no successes

6.6 A bag contains 5 white and 8 red balls. Two drawings of 3 balls are made such that (a) the balls are replaced before the second trial and (b) the balls are not replaced before the second trial. Find the probability that the first drawing will give 3 white and the second, 3 red balls in each case.

[*MD Univ., B.Com.; GND Univ., MA 2001; Kerala Univ., M.Com., 2003*]

6.7 Three groups of workers contain 3 men and one woman, 2 men and 2 women and 1 man and 3 women respectively. One worker is selected at random from each group. What is the probability that the group selected consists of 1 man and 2 women?

[*Nagpur Univ., MCom, 2002*]

6.8 What is the probability that a leap year, selected at random, will contain 53 Sundays?

[*Agra Univ., MCom; Kurukshetra Univ., MCom, 2001; MD Univ., MCom, 2003*]

6.9 A university has to select an examiner from a list of 50 persons, 20 of them women and 30 men, 10 of them knowing Hindi and 40 not, 15 of them being teachers and the remaining 35 not. What is the probability of the university selecting a Hindi-knowing woman teacher? [*Jammu Univ., MCom, 2002*]

Hints and Answers

6.1 (a) P(all heads) = 1/8 (b) P(two heads) = 3/8
(c) P(one head) = 3/8 (d) P(at least one head) = 7/8
(e) P(at least two heads) = 4/8 = 1/2
(f) P(all tails) = 1/8.

6.2 $\text{P(neither a heart nor a king)} = \frac{^{36}C_1}{^{52}C_1} = \frac{36}{52}$

6.3 c(S) = 36; P(total of 11) = 2/36 P(total of 9 or 11) = 7/36

6.4 $\text{P(all graduate)} = \frac{^{5}C_3 \times {}^{15}C_0}{^{20}C_3} = \frac{10 \times 1}{1140} = \frac{1}{114}$

$\text{P(no graduate)} = \frac{^{15}C_3 \times {}^{5}C_0}{^{20}C_2} = \frac{455 \times 1}{1140} = \frac{91}{28}$

$\text{P(at least one graduate)} = 1 - \frac{91}{228} = \frac{137}{228}$

6.5 (a) $\text{P(two successes)} = \frac{13}{25} \times \frac{13}{25} = \frac{169}{625}$

(b) $\text{P(exactly one success)} = \frac{13}{25} \times \frac{12}{25} + \frac{13}{25} \times \frac{12}{25} = \frac{312}{625}$

(c) P(at least one success) = P(exactly one success)

$+ \text{P(two successes)} = \frac{312}{625} + \frac{169}{625} = \frac{481}{625}$

(d) $\text{P(no successes)} = \frac{12}{25} \times \frac{12}{25} = \frac{144}{625}$

6.6 (a) *When balls are replaced:*

Total number of balls in the bag = 5 + 8 = 13.
3 balls can be drawn from 13 in $^{13}C_3$ ways;
3 white balls can be drawn from 5 in $^{5}C_3$ ways;
3 red balls can be drawn from 8 in $^{8}C_3$ ways.

The probability of 3 red balls in the second trial

$$= \frac{^5C_3}{^{13}C_3} = \frac{5}{143}$$

Probability of 3 red balls in the second trial

$$= \frac{^4C_2}{^{12}C_2} = \frac{28}{143}$$

The probability of the compound event

$$\frac{5}{143} \times \frac{28}{143} = \frac{140}{20449} = 0.007$$

(b) *When balls are not replaced:*
At the first trial, 3 white balls can be drawn in $^5C_3$ ways.
The probability of drawing three white balls at the first trial $= \frac{^5C_3}{^{13}C_3} = \frac{5}{143}$

When the white balls have been drawn and not replaced, the bag contains 2 white and 8 red balls. Therefore, at the second trial, 3 balls can be drawn from 10 in $^{10}C_3$ ways and 3 red balls can be drawn from 8 in $^8C_3$ ways.
The probability of 3 red balls in the second trial

$$= \frac{^8C_3}{^{10}C_3} = \frac{7}{15}$$

The probability of the compound event

$$= \frac{5}{142} \times \frac{7}{15} = \frac{7}{429} = 0.016.$$

6.7 There are three possibilities in this case:

(i) Man is selected from the 1st group and women from 2nd and 3rd groups; or
(ii) Man is selected from the 2nd group and women from the 1st and 3rd groups; or
(iii) Man is selected from the 3rd group and women from 1st and 2nd groups.

The probability of selecting a group of 1 man and 2 women is:

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

$$= \frac{9}{32} + \frac{3}{32} + \frac{1}{32} = \frac{13}{32}.$$

6.8 A leap year consists of 366 days, therefore it contains 52 complete weeks and 2 extra days. These 2 days may make the following 7 combinations:

(i) Monday and Tuesday
(ii) Tuesday and Wednesday
(iii) Wednesday and Thursday
(iv) Thursday and Friday
(v) Friday and Saturday
(vi) Saturday and Sunday
(vii) Sunday and Monday

Of these seven equally likely cases, only the last two are favourable. Hence the required probability is 2/7.

6.9 Probability of selecting a woman = 20/50;
Probability of selecting a teacher = 15/50
Probability of selecting a Hindi-knowing candidate =10/50
Since the events are independent, the probability of the university selecting a Hindi-knowing woman teacher = (20/50) × (15/50) × (10/50) = 3/125.

6.5 RULES OF PROBABILITY AND ALGEBRA OF EVENTS

Set theory notations have been used to illustrate the rules of probability and to simplify the calculation of probability of events. In usual notations, the probability of the occurrence of an event A is expressed as:

P(A) = probability of event A occurrence

The probability of a single event is called **marginal (or unconditional) probability.** For example, in the coin tossing example, the marginal probability of a tail or head in a toss can be stated as P(T) or P(H).

Marginal Probability: The unconditional probability of an event occurring.

6.5.1 Rules of Addition

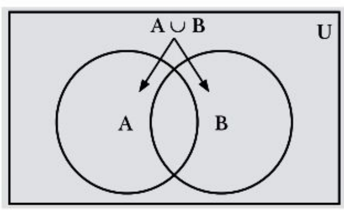
If two or more than two events are likely to occur from a random experiment and we are interested to know the probability of occurrence of at least one of the events, then rules of addition are used to do so.

Mutually Exclusive Events

The rule of addition for mutually exclusive (disjoint), exhaustive and equally likely events states that

If two events A and B are mutually exclusive, exhaustive and equiprobable, then the probability of either event A or B or both occurring is equal to the sum of their individual probabilities.

Figure 6.3
Union of Two Events



This rule is expressed in the following formula:

$$P(A \text{ or } B) = P(A \cup B) = \frac{n(A \cup B)}{n(S)} = \frac{n(A) + n(B)}{n(S)}$$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} = P(A) + P(B) \qquad (6\text{-}1)$$

Venn Diagram: A pictorial representation for showing the sample space and operations involving events. The sample space is represented by a rectangle and events as circles.

where $A \cup B$ (read as 'A *union* B') denotes the union of two events A and B and it is the set of all sample points belonging to A or B or both. This rule can also be illustrated by the **Venn diagram** shown in Fig. 6.3. The two circles contain all the sample points in events A and B. The overlap of the circles indicates that some sample points are contained in both A and B.

Illustration Consider the pattern of arrival of customers at a service counter during the first hour it is open along with its probability:

No. of customers	:	0	1	2	3	4 or more
Probability	:	0.1	0.2	0.3	0.3	0.1

The probability that either 2 or 3 customers will arrive at the service counter during the first hour is determined as

$$P(2 \text{ or } 3) = P(2) + P(3) = 0.3 + 0.3 = 0.6$$

The formula (6-1) can be expanded to include more than two events. In particular, if there are n mutually exclusive events in a sample space, then the probability of the union of these events is given by

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n) \qquad (6\text{-}2)$$

For example, the probability that there may be two or more customers at the service counter during the first hour can be determined by using formula (6-2) as follows:

$$P(2 \text{ or more}) = P(2, 3, 4 \text{ or more}) = P(2) + P(3) + P(4)$$
$$= 0.3 + 0.3 + 0.1 = 0.7$$

A special case of formula (6-1) is for complementary events. Let A be any event and $\bar{A}$ be the complement of A. Then A and $\bar{A}$ are mutually exclusive and exhaustive events. Thus, probability of occurrence of either A or $\bar{A}$ is given by

$$P(A \text{ or } \bar{A}) = P(A) + P(\bar{A}) = P(A) + \{1 - P(A)\} = 1$$

or $$P(A) = 1 - P(\bar{A}) \qquad (6\text{-}3)$$

For example, if a dice is rolled, then the probability whether an odd number of spots occur or do not.

Partially Overlapping (Joint) Events

If events A and B are not mutually exclusive, it is possible for both events to occur simultaneously? This means these events have some sample points in common. Such events are also called *joint* (or *overlapping*) *events*. The sample points in common (belong to both events) represent the joint event $A \cap B$ (read as: A intersection B). The addition rule in this case is stated as

> *If two events A and B are not mutually exclusive, then the probability of either A or B or both occurring is equal to the sum of their individual probabilities minus the probability of A and B occurring together.*

Algebraically, this rule is expressed as

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

or $$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (6\text{-}4)$$

This addition rule can also be illustrated by the Venn diagram shown in Fig. 6.4.

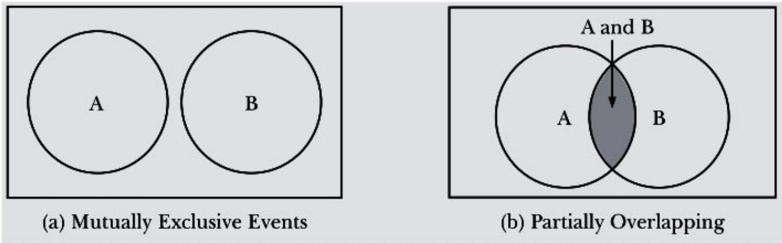


Figure 6.4
Partially Overlapping Events

Illustration Suppose 70 per cent of all tourists who come to India will visit Agra while 60 per cent will visit Goa and 50 per cent of them will visit both Agra and Goa. The probability that a tourist will visit either Goa or Agra or both is obtained by applying formula (6-4) as follows:

$$\begin{aligned} P(\text{Agra or Goa}) &= P(\text{Agra}) + P(\text{Goa}) - P(\text{both Agra and Goa}) \\ &= 0.70 + 0.60 - 0.50 = 0.8 \end{aligned}$$

Consequently, the probability that a tourist will visit neither Agra nor Goa is calculated by

$$P(\text{neither Agra nor Goa}) = 1 - P(\text{Agra or Goa}) = 1 - 0.80 = 0.20$$

The formula (6-4) can be expanded to include more than two events. In particular, if there are three events that are not mutually exclusive, then the probability of the union of these events is given by

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) \\ - P(C \cap A) + P(A \cap B \cap C) \quad (6\text{-}5) \end{aligned}$$

Example 6.4: What is the probability that a randomly chosen card from a deck of cards will be either a king of heart.

Solution: Let event A and B be the king and heart in a deck of 52 cards, respectively. Then, it is given that

Card	*Probability*	*Reason*
King	P(A) = 4/52	4 kings in a deck of 52 cards
Heart	P(B) = 13/52	13 hearts in a deck of 52 cards
King of heart	P(A and (B) = 1/52	1 King of heart in a deck of 52 cards

Using the formula (6-4), we get

$$\begin{aligned} P(A \text{ or } B) &= P(A) + P(B) - P(A \text{ and } B) \\ &= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = 0.3077 \end{aligned}$$

Example 6.5: Of 1000 assembled components, 10 have a working defect and 20 have a structural defect. There is a good reason to assume that no component has both defects. What is the probability that randomly chosen component will have either type of defect?

[Delhi Univ., MBA, 2003]

Solution: Let events A and B be the components which has working defect and has structural defect, respectively. Then it is given that

$$P(A) = 10/1000 = 0.01,\ P(B) = 20/1000 = 0.02 \text{ and } P(A \text{ and } B) = 0$$

The probability that a randomly chosen component will have either type of defect is given by

$$\begin{aligned} P(A \text{ or } B) &= P(A) + P(B) - P(A \text{ and } B) \\ &= 0.01 + 0.02 - 0.0 = 0.03 \end{aligned}$$

Example 6.6: A survey of 200 retail grocery shops revealed following monthly income pattern:

Monthly Income (₹)	*Number of Shops*
Under ₹20,000	102
20,000 to 30,000	61
30,000 and above	37

(a) What is the probability that a particular shop has monthly income under ₹20,000?
(b) What is the probability that a shop selected at random has either an income between ₹20,000 and ₹30,000 or an income of ₹30,000 and more?

Solution: Let the events A, B and C represent the income under three categories, respectively.

(a) Probability that a particular shop has monthly income under ₹20,000 is P(A)=102/200=0.51.
(b) Probability that shop selected at random has income between ₹20,000 and ₹30,000 or ₹30,000 and more is given by

$$P(A \text{ or } B) = P(A) + P(B)$$
$$= \frac{61}{200} + \frac{37}{200} = 0.305 + 0.185 = 0.49$$

Example 6.7: From a sales force of 150 persons, one will be selected to attend a special sales meeting. If 52 of them are unmarried, 72 are college graduates and 3/4 of the 52 that are unmarried are college graduates, find the probability that the salesperson selected at random will be neither single nor a college graduate.

Solution: Let A and B be the events that a salesperson selected is married and that he is a college graduate, respectively. Then, it is given that

$$P(A) = 52/150,\ P(B) = 72/150;\ P(A \text{ and } B) = (3/4)\,(52/150) = 39/150$$

The probability that a salesperson selected at random will be neither single nor a college graduate is:

$$P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B) = 1 - \{P(A) + P(B) - P(A \cap B)\}$$
$$= 1 - \left\{\frac{52}{150} + \frac{72}{150} - \frac{39}{150}\right\} = \frac{13}{30}$$

Example 6.8: From a computer tally based on employer records, the personnel manager of a large manufacturing firm finds that 15 per cent of the firm's employees are supervisors and 25 per cent of the firm's employees are college graduates. He also discovers that 5 per cent are both supervisors and college graduates. Suppose an employee is selected at random from the firm's personnel records, what is the probability of:

(a) selecting a person who is both a college graduate and a supervisor?
(b) selecting a person who is neither a supervisor nor a college graduate?

Solution: Let A and B be the events that the person selected is a supervisor and that he is a college graduate, respectively. Given that

$$P(A) = 15/100;\ P(B) = 25/100;\ P(A \text{ and } B) = 5/100$$

(a) Probability of selecting a person who is both a college graduate and a supervisor is:

$$P(A \text{ and } B) = 5/100 = 0.05$$

(b) Probability of selecting a person who is neither a supervisor nor a college graduate is:

$$P(\bar{A} \text{ and } \bar{B}) = 1 - P(A \text{ or } B) = 1 - [P(A) + P(B) - P(A \text{ and } B)]$$
$$= 1 - \left(\frac{15}{100} + \frac{25}{100} - \frac{5}{100}\right) = \frac{65}{100} = 0.65$$

Example 6.9: The probability that a contractor will get a plumbing contract is 2/3 and the probability that he will not get an electrical contract is 5/9. If the probability of getting at least one contract is 4/5, what is the probability that he will get both?

Solution: Let A and B denote the events that the contractor will get a plumbing and electrical contract, respectively. Given that

$$P(A) = 2/3;\ P(B) = 1 - (5/9) = 4/9;\ P(A \cup B) = 4/5$$
$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{2}{3} + \frac{4}{9} - \frac{4}{5} = \frac{14}{45} = 0.31$$

Thus, the probability that the contractor will get both the contracts is 0.31.

Example 6.10: An MBA applies for a job in two firms X and Y. The probability of his being selected in firm X is 0.7 and being rejected at Y is 0.5. The probability of at least one of his applications being rejected is 0.6. What is the probability that he will be selected by one of the firms?

Solution: Let A and B be the events that an MBA will be selected in firm X and will be rejected in firm Y, respectively. Then, given that

$$P(A) = 0.7,\quad P(\bar{A}) = 1 - 0.7 = 0.3$$

$$P(B) = 0.5,\quad P(\bar{B}) = 1 - 0.5 = 0.5, \text{ and } P(\bar{A} \cup \bar{B}) = 0.6$$

Since $P(A \cap B) = 1 - P(\bar{A} \cup \bar{B}) = 1 - 0.6 = 0.4$

therefore, probability that he will be selected by one of the firms is given by

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.7 + 0.5 - 0.4 = 0.8 \end{aligned}$$

Thus, the probability of an MBA being selected by one of the firms is 0.8.

6.5.2 Rules of Multiplication

Statistically Independent Events

When occurrence of an event does not affect and is not affected by the probability of occurrence of any other event, the event is said to be a *statistically independent event*. There are three types of probabilities under statistical independence: *marginal*, *joint* and *conditional*.

- ***Marginal probability:*** A marginal or unconditional probability is the probability of the occurrence of an event. For example, in the toss of coin, the outcome of each toss is an event that is statistically independent of the outcomes of every other toss of the coin.
- ***Joint probability:*** The probability of two or more independent events occurring together or in succession is called the *joint probability*. The joint probability of two or more independent events is equal to the product of their marginal probabilities. In particular, if A and B are independent events, the probability that both A and B will occur is given by

$$P(AB) = P(A \cap B) = P(A) \times P(B) \qquad (6\text{-}6)$$

  Suppose, in the toss of a coin twice, probability that in both tosses the coin will turn up head is given by

$$P(H_1H_2) = P(H_1) \times P(H_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

  The formula (6-6) is applied because the probability of H or T is not affected by any preceding outcome, i.e. these outcomes are independent.
- ***Conditional probability:*** The *conditional probability* of event A given that event B has already occurred is written as $P(A|B)$. Similarly, we may write $P(B|A)$. The vertical bar is read as 'given' and events appearing to the right of the bar are those that have already occurred. Two events A and B are said to be independent if and only $P(A|B) = P(A)$ or $P(B|A) = P(B)$. Otherwise, events are said to be dependent.

Joint Probability: The probability of two events occurring together or in succession.

Conditional Probability: The probability of an event occurring, given that another event has occurred.

Illustration In the case of independent events the probability of occurrence of either of the events does not depend or affect the occurrence of the other. Thus, probability of a head occurrence in the second toss, given that head resulted in the first toss is still 0.5, i.e., $P(H_2 \mid H_1) = 0.5 = P(H_2)$. It is because probabilities of heads and tails are equal in every toss and do not influence occurrence of head or tail in the previous tosses.

Statistically Dependent Events

When the probability of an event depends upon or affected by the occurrence of any other event, the events are said to be *statistically dependent*. There are three types of probabilities under statistical dependence: *joint*, *conditional* and *marginal*.

Statistical Dependence: The condition when the probability of occurrence of an event is dependent upon, or affected by, the occurrence of some other event.

- ***Joint probability:*** If A and B are dependent events, then their joint probability is no longer equal to the product of their respective probabilities. That is, for dependent events

$$P(A \text{ and } B) = P(A \cap B) \neq P(A) \times P(B)$$

Accordingly, $$P(A) \neq P(A \mid B) \text{ and } P(B) \neq P(B \mid A)$$

The joint probability of events A and B occurring together or in succession under statistical dependences is given by

$$P(A \cap B) = P(A) \times P(B \mid A)$$

or $$P(A \cap B) = P(B) \times P(A \mid B)$$

- ***Conditional probability:*** Under statistical dependence, the conditional probability of event B, given that event A has already occurred is given by

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Similarly, the conditional probability of A, given that event B has occurred, is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

- ***Marginal probability:*** The marginal probability of an event under statistical dependence is the same as the marginal probability of an event under statistical independence.

The marginal probability of events A and B can be written as:

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

Example 6.11: The odds against student X solving a Business Statistics problem are 8 to 6, and odds in favour of student Y solving the problem are 14 to 16.

(a) What is the chance that the problem will be solved if they both try independently of each other?

(b) What is the probability that none of them is able to solve the problem?

[Delhi Univ., MBA, 2007]

Solution: Let A = Event that the first student solves the problem,

B = Event that the second student solves the problem.

$$P(A) = \frac{6}{8+6} = \frac{6}{14} \text{ and } P(B) = \frac{14}{14+16} = \frac{14}{30}$$

(a) Probability that the problem will be solved

= P (at least one of them solves the problem)

= P (A or B) = P(A) + P(B) – P(A and B)

= P(A) + P(B) – P(A) × P(B) [because the events are independent]

$$= \frac{6}{14} + \frac{14}{30} - \frac{6}{14} \times \frac{14}{30} = \frac{73}{105} = 0.695$$

(b) Probability that neither A nor B solves the problem

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

$$= [1 - P(A)] \times [1 - P(B)] = \frac{8}{14} \times \frac{16}{30} = \frac{32}{105} = 0.305$$

Example 6.12: The probability that a new marketing approach will be successful is 0.6. The probability that the expenditure for developing the approach can be kept within the original budget is 0.50. The probability that both of these objectives will be achieved is 0.30. What is the probability that at least one of these objectives will be achieved. For the two events described above, determine whether the events are independent or dependent.

[Delhi Univ., MBA, 2003]

Solution: Let A = Event that the new marketing approach will be successful

B = Event that the expenditure for developing the approach can be kept within the original budget

Given that $P(A) = 0.60$, $P(B) = 0.50$ and $P(A \cap B) = 0.30$

Probability that both events A and B will be achieved is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.60 + 0.50 - 0.30 = 0.80$$

If events A and B are independent, then their joint probability is given by

$$P(A \cap B) = P(A) \times P(B) = 0.60 \times 0.50 = 0.30$$

Since this value is same as given in the problem, events are independent.

Example 6.13: A piece of equipment will function only when the three components A, B and C are working. The probability of A failing during one year is 0.15, that of B failing is 0.05 and that of C failing is 0.10. What is the probability that the equipment will fail before the end of the year?

Solution: Given that

$$P(A \text{ failing}) = 0.15;\ P(A \text{ not failing}) = 1 - P(A) = 0.85$$
$$P(B \text{ failing}) = 0.05;\ P(B \text{ not failing}) = 1 - P(B) = 0.95$$
$$P(C \text{ failing}) = 0.10;\ P(C \text{ not failing}) = 1 - P(C) = 0.90$$

Since all the three events are independent, therefore the probability that the equipment will work is given by

$$P(\bar{A} \cap \bar{B} \cap \bar{B}) = P(\bar{A}) \times P(\bar{B}) \times P(\bar{C})$$
$$= 0.85 \times 0.95 \times 0.90 = 0.726$$

Probability that the equipment will fail before the end of the year is given by

$$P(A \cup B \cup C) = 1 - P(\bar{A} \cap \bar{B} \cap \bar{C})$$
$$= 1 - P(\bar{A}).\ P(\bar{B}).\ P(\bar{C})$$
$$= 1 - \{0.85 \times 0.95 \times 0.90\} = 1 - 0.726 = 0.274$$

Example 6.14: A market research firm is interested in surveying certain attitudes in a small community. There are 125 households broken down according to income, ownership of a telephone and ownership of a TV.

	Households with Annual Income of ₹80,000 or Less		*Households with Annual Income Above ₹80,000*		*Total*
	Telephone Subscriber	*No Telephone*	*Telephone Subscriber*	*No Telephone*	
Own TV set	27	20	18	10	75
No TV set	18	10	12	10	50
Total	45	30	30	20	125

(a) What is the probability of getting a TV owner in a random draw?

(b) If a household has an income of over ₹80,000 and is a telephone subscriber, what is the probability that he owns a TV?

(c) What is the conditional probability of drawing a household that owns a TV, given that the household is a telephone subscriber?

(d) Are the events 'ownership of a TV' and 'telephone subscriber' statistically independent? Comment. *[Himachal Univ., MBA, 2006]*

Solution: (a) Probability of drawing a TV owner at random, P(TV owner) = 75/125 = 0.6

(b) There are 30(18 + 12) people whose household income is above ₹80,000 and are also telephone subscribers. Out of these, 18 own TV sets. Hence, the probability of this group of people having a TV set is : 18/30 = 0.6.

(c) Out of 75(27 + 18 + 18 + 12) households who are telephone subscribers, 45(27 + 18) households have TV sets. Hence, the conditional probability of drawing a household that owns a TV given that the household is a telephone subscriber is: 45/75 = 0.6.

(d) Let A and B be the events representing TV owners and telephone subscribers respectively. The probability of a person owning a TV, P(A)= 75/125. The probability of a person being a telephone subscriber, P(B) = 75/125.

The probability of a person being a telephone subscriber as well as a TV owner is:

$$P(A \text{ and } B) = 45/125 = 9/25$$

But $P(A) \times P(B) = (75/125)\ (75/125) = 9/25$

Since $P(AB) = P(A) \times P(B)$, therefore, we conclude that the events 'ownership of a TV' and 'telephone subscriber' are statistically independent.

Example 6.15: A company has two plants to manufacture scooters. Plant I manufactures 80 per cent of the scooters and Plant II manufactures 20 per cent. In plant I, only 85 out of 100 scooters are considered to be of standard quality. In plant II, only 65 out of 100 scooters are considered to be of standard quality. What is the probability that a scooter selected at random came from plant I, if it is known that it is of standard quality?

[*Madras Univ., M.Com., 2006; Delhi Univ., MBA, 2008*]

Solution: Let A = Scooter purchased is of standard quality
B = Scooter is of standard quality and came from plant I
C = Scooter is of standard quality and came from plant II
D = Scooter came from plant I

The percentage of scooters manufactured in plant I that are of standard quality is 85 per cent of 80 per cent, that is, $0.85 \times (80 \div 100) = 68$ per cent or $P(B) = 0.68$.

The percentage of scooters manufactured in plant II that are of standard quality is 65 per cent of 20 per cent, that is, $0.65 \times (20 \div 100) = 13$ per cent or $P(C) = 0.13$.

The probability that a customer obtains a standard quality scooter from the company is $0.68 + 0.13 = 0.81$.

The probability that the scooters selected at random came from plant I, if it is known that it is of standard quality is given by

$$P(D|A) = \frac{P(D \text{ and } A)}{P(A)} = \frac{0.68}{0.81} = 0.84$$

Example 6.16: A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is 1/7 and that of wife's selection is 1/5. What is the probability that

(a) both of them will be selected.
(b) only one of them will be selected.
(c) none of them will be selected. [*Bharthidasan Univ., M.Com., 2006; Delhi Univ., MBA, 2009*]

Solution: Let A and B be the events of the husband's and wife's selection, respectively. Given that $P(A) = 1/7$ and $P(B) = 1/5$.

(a) The probability that both of them will be selected is

$$P(A \text{ and } B) = P(A)\,P(B) = (1/7) \times (1/5) = 1/35 = 0.029$$

(b) The probability that only one of them will be selected is

$$\begin{aligned} P[(A \text{ and } \bar{B}) \text{ or } (B \text{ and } \bar{A})] &= P(A \text{ and } \bar{B}) + P(B \text{ and } \bar{A}) \\ &= P(A)\,P(\bar{B}) + P(B)\,P(\bar{A}) \\ &= P(A)\,[1 - P(B)] + P(B)\,[1 - P(A)] \\ &= \frac{1}{7}\left(1 - \frac{1}{5}\right) + \frac{1}{5}\left(1 - \frac{1}{7}\right) = \left(\frac{1}{7} \times \frac{4}{5}\right) + \left(\frac{1}{5} \times \frac{6}{7}\right) \\ &= \frac{10}{35} = 0.286 \end{aligned}$$

(c) The probability that none of them will be selected is

$$P(\bar{A}) \times P(\bar{B}) = (6/7) \times (4/5) = 24/35 = 0.686$$

Example 6.17: The odds that A speaks the truth is 3:2 and the odds that B speaks the truth is 5 : 3. In what percentage of cases are they likely to contradict each other on an identical point? [*Delhi Univ., MBA, 2009*]

Solution: Let X and Y denote the events that A and B speak truth, respectively. Given that

$$P(X) = 3/5;\quad P(\bar{X}) = 2/5;\quad P(Y) = 5/8;\quad P(\bar{Y}) = 3/8$$

The probability that A speaks the truth and B speaks a lie is (3/5) (3/8) = 9/40

The probability that B speaks the truth and A speaks a lie is (5/8) (2/5) = 10/40. So, the compound probability is $\frac{9}{40}+\frac{10}{40}=\frac{19}{40}$

Hence, percentage of cases in which they contradict each other is (19/40) × 100 = 47.5 per cent.

Example 6.18: A man wants to marry a girl having qualities: (i) white complexion (WC), probability of getting such a girl is one is twenty; (ii) handsome dowry(HD), the probability of getting such a girl is one in fifty; (iii) westernised manners (WM), and etiquettes, the probability of getting such a girl is one in hundred. Find out the probability of his getting married to such a girl when the possession of these three attributes is independent.

[Punjab Univ., B. Com., 2000; Kashmir Univ., M.Com., 2002]

Solution: Probability of a girl with white complexion, 1/20 = 0.5

Probability of a girl with handsome dowry, 1/50 = 0.02

Probability of a girl with westernised manners, 1/100 = 0.01

Since the events are independent, the probability of simultaneous occurrence of all these qualities is:

$$P(WC \cap HD \cap WM) = P(WC) \times P(HD) \times P(WM)$$
$$= 0.05 \times 0.02 \times 0.01 = 0.0001.$$

Example 6.19: A problem of statistics is given to two students *A* and *B*. The odds in favour of A solving the problem are 6 to 9 and against B solving the problem 12 to 10. If A and B attempt to solve the problem independently, then find the probability of the problem being solved. *[Delhi Univ., B.Com.(H), 2000]*

Solution: Let A = problem can be solved by A;

B = problem can be solved by B.

Then based on the data given in the problem, we have

$$P(A) = \frac{6}{6+9}=\frac{6}{15}; \qquad P(\bar{A}) = 1-\frac{6}{15}=\frac{9}{15}$$

$$P(B) = 1-\frac{12}{12+10}=\frac{5}{11}; \quad P(\bar{B}) = 1-\frac{5}{11}=\frac{6}{11}.$$

$$P(\bar{A} \text{ and } \bar{B}) = P(\bar{A}) \times (\bar{B}) = \frac{9}{15}\times\frac{6}{11}=\frac{18}{55}$$

Probability that the problem being solved is:

$$P(A \text{ or } B) = 1-\frac{18}{55}=\frac{37}{55}=0.673.$$

Example 6.20: A can solve 90 per cent of the problems given in a book and B can solve 70 per cent. What is the porbability that at least one of them will solve a problem selected at random? *[Kurukshetra, Univ., B.Com., 2000; M.D. Univ., M.Com., 2004]*

Solution: Probability that A will not be able to solve the problem

$$P(\bar{A}) = 1-\frac{9}{10}=\frac{1}{10}.$$

Probability that *B* will not be able to solve the problem

$$P(\bar{B}) = 1-\frac{7}{10}=\frac{3}{10}.$$

Probability that none of them will be able to solve the problem

$$P(\bar{A} \text{ and } \bar{B}) = P(\bar{A} \cap \bar{B}) = \frac{1}{10}\times\frac{3}{10}=\frac{3}{100}.$$

Hence, probability that at least one of them will solve the problem

$$P(\text{A or B}) = P(\text{A} \cup \text{B}) = 1 - \frac{3}{100} = \frac{97}{100}.$$

Example 6.21: One bag contains 4 white and 2 black balls. Another contains 3 white and 5 black balls. If one ball is drawn from each bag, find the probability that (a) both are white, (b) both are black, and (c) one is white and one is black. *[Madras Univ., B.Com., 2005]*

Solution: (a) Probability of drawing a white ball from the first bag, $P(W_1) = 4/6$.

Probability of drawing a white ball from the second bag, $P(W_2) = 3/8$.

Since the events are independent, the probability that both the balls are white,

$$P(W_1 \text{ and } W_2) = \frac{4}{6} \times \frac{3}{8} = \frac{1}{4}.$$

(b) Probability of drawing a black ball from the first bag, $P(B_1) = 2/6$.

Probability of drawing a black ball from the second bag, $P(B_2) = 5/8$.

Probability that both the balls are black, $P(B_1 \text{ and } B_2) = \frac{2}{6} \times \frac{5}{8} = \frac{5}{24}$.

(c) That event 'one is white one is black' is the same as saying that 'either first ball is white and second is black or first ball is black the second is white.' Hence,

Probability that one is white and one is black

$$P(\text{W and B}) = P(\text{W}) \times P(\bar{\text{B}}) + P(\bar{\text{W}})P(\text{B}) = \left(\frac{4}{6}\right)\left(\frac{5}{8}\right) + \left(\frac{2}{6}\right)\left(\frac{3}{8}\right) = \frac{13}{24}.$$

Example 6.22: A bag contains 5 white and 3 black balls. Two balls are drawn at random one after the other without replacement. Find the probability that both balls drawn are black. *[Bhartidasan Univ., 2000]*

Solution: Probability of drawing a black ball in the first attempt is

$$P(B_1) = \frac{3}{5+3} = \frac{3}{8}.$$

Probability of drawing the second ball given that the first ball drawn is black

$$P(B_2 | B_1) = \frac{P(B_1 \cap B_2)}{P(B_1)} = \frac{(3/8) \times (2/7)}{3/8} = \frac{2}{7}.$$

The probability that both balls drawn are black is given by

$$P(B_1 | B_2) = P(B_1) \times P(B_2 | B_1) = \frac{3}{8} \times \frac{2}{7} = \frac{3}{28}.$$

Example 6.23: In a certain town, male and female form 50 per cent of the population. It is known that 20 per cent of the male and 5 per cent of the female are unemployed. A research student is studying the employment person at random. What is the probability that the person selected is (a) male (b) female?

[Delhi Univ., M.Com., 2000; Kumaon Univ., MBA, 2005]

Solution: The data of the problem according to the sex and employment status is summarized as follows:

	Employed	*Unemployed*	*Total*
Male	0.40	0.10	0.50
Female	0.475	0.025	0.50
Total	0.875	0.125	1.00

Let M and F be a male and a female chosen for study, respectively, where U denote male and female chosen is unemployed.

(a) $P(M | U) = \frac{P(M \cap U)}{P(U)} = \frac{0.10}{0.125} = 0.80$.

(b) $P(F | U) = \frac{P(F \cap U)}{P(U)} = \frac{0.025}{0.125} = 0.20$.

Example 6.24: The data for the promotion and academic qualification of a company is given below:

Promotional Status	*Academic Qualification*		
	MBA	*Non-MBA*	*Total*
Promoted	0.14	0.26	0.40
Non-promoted	0.21	0.39	0.60
Total	0.35	0.65	1.00

(a) Calculate the conditional probability of promotion after an MBA has been identified.

(b) Calculate the conditional probability that it is an MBA when a promoted employee has been chosen.

(c) Find the probability that a promoted employee was an MBA. *[IGNOU, 2001]*

Solution: Let A and B be an event denoting MBA qualification and promotion, respectively. Then given that $P(A) = 0.35$, $P(\bar{A}) = 0.65$, $P(B) = 0.40$, $P(\bar{B}) = 0.60$ and $P(A \cap B) = 0.14$

(a) The probability of being 'promoted' after an MBA employee has been identified is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.14}{0.35} = 0.40$$

(b) If a promoted employee has been chosen, then the probability that the person is an MBA is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.14}{0.40} = 0.35$$

(c) The probability that a promoted employee was an MBA is:

$$P(A \cap B) = P(A) \times P(B|A) = 0.35 \times 0.40 = 0.14$$

or $$= P(B) \times P(A|B) = 0.40 \times 0.35 = 0.14$$

Example 6.25: The probability that a trainee will remain with a company is 0.6. The probability that an employee earns more than ₹30,000 per month is 0.5. The probability that an employee who is a trainee remained with the company or who earns more than ₹30,000 per month is 0.7. What is the probability that an employee earns more than ₹30,000 per month given that he is a trainee who stayed with the company?

Solution: Let A and B be the events that a trainee who remained with the company and the event that an employee earns more than ₹30,000, respectively. Given that

$$P(A) = 0.6,\ P(B) = 0.5 \text{ and } P(A \text{ or } B) = P(A \cup B) = 0.7$$

The probability that an employee earns more than ₹30,000, given that he is trainee who remained with the company, is given by

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

We know that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

or $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.6 + 0.5 - 0.7 = 0.4$

Hence, the required probability is

$$P(B|A) = \frac{(A \cap B)}{P(A)} = \frac{0.4}{0.6} = 0.667$$

Example 6.26: Two computers A and B are to be marketed. A salesman who is assigned the job of finding customers for them has 60 per cent and 40 per cent chances of succeeding for computers A and B, respectively. The two computers can be sold independently. Given that he was able to sell at least one computer, what is the probability that computer A has been sold? *[IGNOU, MBA, 2002; Delhi Univ., MBA, 2002, 2005]*

Solution: Let E_1 and E_2 be the events that computer A and B is marketed, respectively. It is given that $P(E_1) = 0.60$, $P(E_2) = 0.40$. Thus

$$P(E_1 \text{ and } E_2) \text{ or } P(E_1 \cap E_2) = P(E_1).\ P(E_2) = 0.60 \times 0.40 = 0.24$$

Hence, the probability that computer A has been sold given that the salesman was able to sell at least one computer is given by

$$P(E_1 \mid E_1 \cup E_2) = \frac{P\{E_1 \cap (E_1 \cup E_2)\}}{P(E_1 \cup E_2)} = \frac{P(E_1)}{P(E_1 \cup E_2)}$$

$$= \frac{P(E_1)}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} = \frac{0.60}{0.60 + 0.40 - 0.24} = \frac{0.60}{0.76} = 0.789$$

Example 6.27: A study of job satisfaction was conducted for four occupations: Cabin-maker, Lawyer, Doctor and Systems Analyst. Job satisfaction was measured on a scale of 0–100. The data obtained are summarized in the following table:

Occupation	*Under 50*	*50–59*	*60–69*	*70–79*	*80–89*	*Total*
Cabin-maker	0	2	4	3	1	10
Lawyer	6	2	1	1	0	10
Doctor	0	5	2	1	2	10
Systems Analyst	2	1	4	3	0	10
	8	10	11	8	3	40

(a) Develop a joint probability table.
(b) What is the probability of one of the participants studied had a satisfaction score in 80's?
(c) What is the probability of a satisfaction score in the 80's, given the study participant was a doctor?
(d) What is the probability of one of the participants studied was a lawyer.
(e) What is the probability of one of the participants was a lawyer and received a score under 50?
(f) What is the probability of a satisfaction score under 50 given a person is a lawyer.
(g) What is the probability of a satisfaction score of 70 or higher? *[Delhi Univ., MBA, 2006]*

Solution: (a) Joint probability table is given below:

Occupation	*Under 50*	*50–59*	*60–69*	*70–79*	*80–89*
Cabin-maker	0.000	0.050	0.100	0.075	0.250
Lawyer	0.150	0.050	0.025	0.025	0.250
Doctor	0.000	0.125	0.050	0.025	0.250
Systems Analyst	0.050	0.025	0.100	0.075	0.250

(b) P(Satisfaction score in the 80's) = 3/40

(c) P(Satisfaction score in 80's, given participant was doctor) $= \frac{2/40}{10/40} = \frac{1}{5}$

(d) P(Participant was doctor) = 10/40

(e) P(Lawyer and score under 50) $= \frac{P(\text{Lawyer} \cap \text{Score under 50})}{P(\text{Score under 50})} = \frac{6}{40}$

(f) P(Score under 50 Lawyer) $= \frac{P(\text{Score under 50} \cap \text{Lawyer})}{P(\text{Lawyer})} = \frac{6/40}{10/40} = \frac{6}{10}$

(g) P(Satisfaction score of 70 or higher) = P(Score of 70 and above)
+ P(Score of 80 and above) $= \frac{8}{40} + \frac{3}{40} = \frac{11}{40}$

Example 6.28: A market survey was conducted in four cities to find out the preference for brand A soap. The responses are shown below:

	Delhi	*Kolkata*	*Chennai*	*Mumbai*
Yes	45	55	60	50
No	35	45	35	45
No opinion	5	5	5	5

(a) What is the probability that a consumer selected at random, preferred brand A?
(b) What is the probability that a consumer preferred brand A and was from Chennai?

(c) What is the probability that a consumer preferred brand A, given that he was from Chennai?

(d) Given that a consumer preferred brand A, what is the probability that he was from Mumbai? *[Delhi Univ., MBA, 2002; Kumaon Univ., MBA, 1999]*

Solution: The information from responses during market survey is as follows:

	Delhi	*Kolkata*	*Chennai*	*Mumbai*	*Total*
Yes	45	55	60	50	210
No	35	45	35	45	160
No opinion	5	5	5	5	20
Total	85	105	100	100	390

If X is an event that a consumer selected at random preferred brand A, find

(a) The probability that a consumer selected at random preferred brand A is

$$P(X) = 210/390 = 0.5398$$

(b) The probability that a consumer preferred brand A and was from Chennai (C) is

$$P(X \cap C) = 60/390 = 0.1538$$

(c) The probability that a consumer preferred brand A, given that he was from Chennai:

$$P(X|C) = \frac{P(A \cap C)}{P(C)} = \frac{60/390}{100/390} = \frac{0.153}{0.256} = 0.597$$

(d) The probability that the consumer belongs to Mumbai, given that he preferred brand A

$$P(M|X) = \frac{P(M \cap X)}{P(X)} = \frac{50/390}{210/390} = \frac{0.128}{0.538} = 0.237$$

Example 6.29: The personnel department of a company has records which show the following analysis of its 200 engineers.

Age	*Bachelor's Degree Only*	*Master's Degree*	*Total*
Under 30	90	10	100
30 to 40	20	30	50
Over 40	40	10	50
Total	150	50	200

If one engineer is selected at random from the company, find

(a) The probability that he has only a bachelor's degree.
(b) The probability that he has a master's degree, given that he is over 40.
(c) The probability that he is under 30, given that he has only a bachelor's degree.

[Kumaon Univ., MBA, 2004]

Solution: Let A, B, C and D denote the events that an engineer in under 30 years of age, 40 year of age, has bachelor's degree only and has a master's degree, respectively. Therefore,

(a) The probability of an engineer who has only a bachelor's degree is

$$P(C) = 150/200 = 0.75$$

(b) The probability of an engineer who has a master's degree, given that he is over 40 years is

$$P(D|B) = \frac{P(D \cap B)}{P(B)} = \frac{10/200}{50/200} = \frac{10}{50} = 0.20$$

(c) The probability of an engineer who is under 30 years, given that he has only bachelor's degree is

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{90/200}{150/200} = \frac{90}{150} = 0.60$$

Self-practice Problems 6B

6.10 Mr. X has 2 shares in a lottery in which there are 2 prizes and 5 blanks. Mr. Y has 1 share in a lottery in which there is 1 prize and 2 blanks. Show that the chance of Mr. X's success to that of Mr. Y's is 15:7.

6.11 Explain whether or not each of the following claims could be correct:

(a) A businessman claims that the probability that he will get contract A is 0.15 and that he will get contract B is 0.20. Furthermore, he claims that the probability of getting A or B is 0.50.

(b) A market analyst claims that the probability of selling ten million rupees of plastic A or five million rupees of plastic B is 0.60. He also claims that the probability of selling ten million rupees of A and five million rupees of B is 0.45.

6.12 The probability that an applicant for a Management Accountant's job has a postgraduate degree is 0.3, he has had some work experience as a Chief Financial Accountant is 0.7, and that he has both is 0.2. Out of 300 applicants, approximately, what number would have either a postgraduate degree or some professional work experience?

6.13 A can hit a target 3 times in 5 shots; B, 2 times in 5 shots; C, 3 times in 4 shots. They fire a volley. What is the probability that 2 shots hit?

6.14 A problem in business statistics is given to five students, A, B, C, D and E. Their chances of solving it are 1/2, 1/3, 1/4, 1/5 and 1/6 respectively. What is the probability that the problem will be solved?

[*Madras Univ., B.Com., 1996; Kumaon Univ., MBA, 2000*]

6.15 Three persons A, B and C are being considered for appointment as Vice-Chancellor of a university, and whose chances of being selected for the post are in the proportion 14:2:3 respectively. The probability that A if selected, will introduce democratization in the university structure is 0.3, and the corresponding probabilities for B and C doing the same are respectively 0.5 and 0.8. What is the probability that democratization would be introduced in the university?

6.16 There are three brands, say X, Y and Z of an item available in the market. A consumer chooses exactly one of them for his use. He never buys two or more brands simultaneously. The probabilities that he buys brands X, Y and Z are 0.20, 0.16 and 0.45, respectively.

(a) What is the probability that he does not buy any of the brands?

(b) Given that a customer buys some brand, what is the probability that he buys brand X?

6.17 There is 50-50 chance that a contractor's firm, A, will bid for the construction of a multi-storied building. Another firm, B, submits a bid and the probability is 3/5 that it will get the job, provided that firm A does not submit a bid. If firm A submits a bid, the probability that firm B will get the job is only 2/3. What is the probability that firm B will get the job?

6.18 Plant I of XYZ manufacturing organization employs 5 production and 3 maintenance foremen, plant II of same organization employs 4 production and 5 maintenance foremen. From any one of these plants, a single selection of two foremen is made. Find the probability that one of them would be a production and the other a maintenance foreman.

[*Bombay Univ., MMS, 2001*]

6.19 If a machine is correctly set up, it will produce 90 per cent acceptable items. If it is incorrectly setup, it will produce 40 per cent acceptable items. Past experience shows that 80 per cent of the setups are correctly done. If after a certain setup, the machine produces 2 acceptable items as the first 2 pieces. Find the probability that the machine is correctly set up?

6.20 A firm plans to bid ₹300 per tonne for a contract to supply 1,000 tonnes of a metal. It has two competitors A and B. It assumes the probability of A bidding less than ₹300 per tonne to be 0.3 and B's bid to be less than ₹300 per tonne to be 0.7. If the lowest bidder gets all the business and the firms bid independently, what is the expected value of the contract to the firm?

6.21 An investment consultant predicts that the odds against the price of a certain stock going up during the next week are 2 : 1 and odds in favour of the price remaining the same are 1:3. What is the probability that the price of the stock will go down during the next week?

6.22 An article manufactured by a company consists of two parts A and B. In the process of manufacture of part A, 9 out of 100 are likely to be defective. Similarly, 5 out of 100 are likely to be defective in the manufacture of part B. Calculate the probability that the assembled part will not be defective.

6.23 A product is assembled from three components X, Y and Z, the probability of these components being defective is 0.01, 0.02 and 0.05, respectively. What is the probability that the assembled product will not be defective?

6.24 The daily production of a machine producing a very complicated item gives the following probabilities for the number of items produced: P(1) = 0.20, P(2) = 0.35 and P(3) = 0.45. Furthermore, the probability of defective items being produced is 0.02. Defective items are assumed to occur independently. Determine the probability of no defectives during a day's production.

6.25 The personnel department of a company has records which show the following analysis of its 200 engineers:

Age (Years)	*Bachelor's Degree only*	*Master's Degree*	*Total*
Under 30	90	10	100
30 to 40	20	30	50
Over 40	40	10	50
	150	50	200

If one engineer is selected at random from the company, find

(a) the probability that he has only a bachelor's degree.

(b) the probability that he has a master's degree given that he is over 40.

(c) the probability that he is under 30 given that he has only a bachelor's degree.

[*HP Univ., MBA; Kumaon Univ., MBA 2005*]

6.26 In a certain town, males and females form 50 per cent of the population. It is known that 20 per cent of the males and 5 per cent of the females are unemployed. A research student studying the employment situation selects unemployed persons at random. What is the probability that the person selected is (a) male, (b) female? [*Delhi Univ. M.Com., 2002; Kumaon Univ., MBA, 2001*]

6.27 You note that your officer is happy in 60 per cent cases of your calls. You have also noticed that if he is happy, he accedes to your requests with a probability of 0.4, whereas if he is not happy, he accedes to your requests with a probability of 0.1. You call on him one day and he accedes to your request. What is the probability of his being happy? [*HP, MBA, 2005*]

6.28 In a telephone survey of 1000 adults, respondents were asked about the expenses on a management education and the relative necessity of some form of financial assistance. The respondents were classified according to whether they currently had a child studying in a school of management and whether they thought that the loan burden for most management students is: too high, right amount, or too little. The proportions responding in each category are given below.

	Too High (A)	*Right Amount (B)*	*Too Little (C)*
Child studying management (D) :	0.35	0.08	0.01
No child studying management (E) :	0.25	0.20	0.11

Suppose one respondent is chosen at random from this group. Then

(a) What is the probability that the respondent has a child studying management.

(b) Given that the respondent has a child studying management, what is the probability that he/she ranks the loan burden as 'too high'.

(c) Are events D and A independent? Explain.

6.29 In a colour preference experiment, eight toys are placed in a container. The toys are identical except for colour—two are red, and six are green. A child is asked to choose two toys at random. What is the probability that the child chooses the two red toys?

6.30 A survey of executives dealt with their loyalty to the company. One of the questions was, 'If you were given an offer by another company equal to or slightly better than your present position, would you remain with the company?' The responses of 200 executives in the survey cross-classified with their length of service with the company are shown below:

	Length of Service				
Loyalty	*Less than 1 year*	*1–5 years*	*6–10 years*	*More than 10 years*	*Total*
Would remain :	10	30	5	75	120
Would not remain :	25	15	10	30	80

What is the probability of randomly selecting an executive who is loyal to the company (would remain) and who has more than 10 years of service.

Hints and Answers

6.10 Considering Mr X's chances of success.

A = Event that 1 share brings a prize and 1 share goes blank.

B = Event that both the shares bring prizes.

C = event that X succeeds in getting at least one prize = A ∪ B.

Since A and B are mutually exclusive, therefore

$$P(C) = P(A \cup B) = P(A) + P(B) = \frac{{}^2C_1 \times {}^5C_1}{{}^7C_2} + \frac{{}^2C_2 \times {}^5C_0}{{}^7C_2}$$

If D is an event that Y succeeds in getting a prize, then

$$P(D) = \frac{{}^1C_1}{{}^3C_1} = \frac{1}{3}$$

$$\text{X's chance of success: Y's chance of success} = \frac{15}{21} : \frac{1}{3} = 15 : 7.$$

6.11 (a) $P(A \cap B) = -0.15$

(b) $P(A) + P(B) = 1.05$

6.12 Let A = Applicant has P.G degree;

B = Applicant has work experience;

Given, P(A) = 0.3, P(B) = 0.7 and $P(A \cap B) = 0.2$.

$300 \times P(A \cup B) = 300[P(A) + P(B) - P(A \cap B)] = 240$

6.13 The required event that two shots may hit the target can happen in the following mutually exclusive cases:

(i) A and B hit and C fails to hit the target.

(ii) A and C hit and B fails to hit the target.

(iii) B and C hit and A fails to hit the target.

Hence, the required probability that any two shots hit is given by, P = P(i) + P(ii) + P(iii).

Let E_1, E_2 and E_3 be the event of hitting the target by A, B and C respectively. Therefore

$$P(i) = P(E_1 \cap E_2 \cap \bar{E}_3) = P(E_1)\,P(E_2)\,P(\bar{E}_3)$$

$$= \left(\frac{3}{5}\right)\left(\frac{2}{5}\right)\left(1-\frac{3}{4}\right) = \frac{6}{100}$$

$$P(ii) = P(E_1 \cap \bar{E}_2 \cap E_3) = \left(\frac{3}{5}\right)\left(1-\frac{2}{5}\right)\left(\frac{3}{4}\right) = \frac{27}{100}$$

$$P(iii) = P(\bar{E}_1 \cap E_2 \cap E_3) = \left(1-\frac{3}{5}\right)\left(\frac{2}{5}\right)\left(\frac{3}{4}\right) = \frac{12}{100}$$

Since all the three events are mutually exclusive events, hence the required probability is given by

$$P(i) + P(ii) + P(iii) = \frac{6}{100} + \frac{27}{100} + \frac{12}{100} = \frac{9}{20}$$

6.14 P(problem will be solved)

= 1 – P(problem is not solved)

= 1 – P(all students fail to solve the problem)

$= 1 - P(\bar{A} \cap \bar{B} \cap \bar{C} \cap \bar{D} \cap \bar{E})$

$= 1 - P(\bar{A})\,P(\bar{B})\,P(\bar{C})\,P(\bar{D})\,P(\bar{E})$

$$= 1 - \left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{4}\right)\left(1-\frac{1}{5}\right)\left(1-\frac{1}{6}\right)$$

$$= 1 - \frac{1}{6} = \frac{5}{6}$$

6.15 P(D) = Prob. of the event that democratization would be introduced

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

$= P[(A \cap D) + P(B \cap D) + P(C \cap D)$

$= P(A)\,P(D \mid A) + P(B)\,P(D \mid B) + P(C)P(D \mid C)$

$$= 0.3\left(\frac{4}{9}\right) + 0.5\left(\frac{2}{9}\right) + 0.8\left(\frac{3}{9}\right) = 0.51$$

6.16 (a) P(Customer does not buy any brand)

$= P[\bar{X} \cap \bar{Y} \cap \bar{Z}] = 1 - P[(X \cup Y \cup Z]$

$= 1 - [P(X) + P(Y) + P(Z)]$

$= 1 - [0.20 + 0.16 + 0.45] = 0.19$

(b) P(Customer buys brand X) = $P[X \mid (X \cup Y \cup Z)]$

$$= \frac{P[X \cap (X \cup Y \cup Z)]}{P(X \cup Y \cup Z)}$$

$$= \frac{P(X)}{P(X)+P(Y)+P(Z)-P(X\cap Y)-P(Y\cap Z)-P(X\cap Z)+P(X\cap Y\cap Z)}$$

$$= \frac{0.2}{0.2+0.16+0.45-0-0-0-0} = 0.247$$

6.17 P(A) = 1/2, P(B | A) = 2/3 and $(B \mid \bar{A}) = 3/5$.

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

$= P(A) \times P(B|A) + P(\bar{A}) \cdot P(B \mid \bar{A})$

$$= \frac{1}{2}\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{3}{5} = \frac{19}{30}.$$

6.18 Let E_1 and E_2 be the events that Plant I and II is selected respectively. Then, probability of the event E that in a batch of 2, one is the production and the other is the maintenance man is

$P(E) = P(E_1 \cap E) + P(E_2 \cap E)$

$= P(E_1)\;P(E|E_1) + P(E_2)\,P(E|E_2)$

$$= \frac{1}{2}\cdot\frac{{}^5C_1 \cdot {}^3C_1}{{}^8C_2} + \frac{1}{2}\cdot\frac{{}^4C_1 \cdot {}^5C_1}{{}^9C_2}$$

$$= \frac{1}{2}\cdot\frac{15}{28} + \frac{1}{2}\cdot\frac{5}{9} = \frac{275}{504}$$

6.19 Let A = event that the item is acceptable;

B_1 and B_2 = events that machine is correctly and incorrectly setup, respectively.

Given, $P(A|B_1) = 0.9$; $P(A \mid B_2) = 0.4$; $P(B_1) = 0.8$ and $P(B_2) = 0.2$. Then $P(B_1 \mid A) = 0.9$.

6.20 There are two competitors A and B and the lowest bidder gets the contract.

Value of plan = 300 × 1,000 = 3,00,000

Contractor A : P(Bid < 300) = 0.3;
P(Bid ≥ 300) = 0.7

Contractor B : P(Bid < 300) = 0.7;
P(Bid ≥ 300) = 0.3

(i) If both bids are less than ₹300, probability is 0.3 × 0.7 = 0.21. Therefore plan value is:

3,00,000 × 0.21 = 63,000.

(ii) If A bids less than 300 and B bids more than 300, probability is 0.3 × 0.3 = 0.9. Therefore, plan value is: 3,00,000 × 0.09 = 27,000.

(iii) B bids less than 300 while A bids more than 300, probability is: 0.7 × 0.7 = 0.49. Therefore plan value is: 3,00,000 × 0.49 = 1,47,000.

Therefore, expected value of plan is

63,000 + 27,000 + 1,47,000 = 2,37,000.

6.21 P(Price of a certain stock not going up) = 2/3

P(Price of a certain stock remaining same) = 1/4

The probability that the price of the stock will go down during the next week

= P(Price of the stock not going up and not remaining same)

= P(Price of the stock not going up × P(price of the stock not remaining same)

$$= \left(\frac{2}{3}\right)\times\left(1-\frac{1}{4}\right)=\left(\frac{2}{3}\right)\times\left(\frac{3}{4}\right)=\frac{1}{2} = 0.5$$

6.22 The assembled part will be defective if any of the parts is defective.

The probability of the assembled part being defective:

= P[Any of the part is defective]

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

$$= \frac{9}{100}+\frac{5}{100}-\left(\frac{9}{100}\right)\times\left(\frac{5}{100}\right) = 0.1355$$

The probability that assembled part is not defective

= 1 – 0.1355 = 0.8645.

6.23 Let A, B and C denote the respective probabilities of components X, Y and Z being defective.

P(A) = 0.01, P(B) = 0.02, P(C) = 0.05

P(A or B or C) = P(A) + P(B) + P(C) – P(AB) – P(BC) – P(AC) + P(ABC)

= 0.01 + 0.02 + 0.05 – 0.0002 – 0.0010 – 0.0005 + 0.00001 = 0.0784

Hence the probability that the assembled product will not be defective = 1 – 0.0784 or 0.9216.

6.24 Let A be the event that no defective item is produced during a day. Then

P(A) = P(I)P(A|I) + P(2)P(A|2) + P(3)P(A|3)

The probability that a defective item is produced = 0.02. Probability that a non-defective item is produced = 1 – 0.02 = 0.98. Also defectives are assumed to occur independently, therefore

P(A | I) = 0.98, P(A | 2) = (0.98) (0.98) and

P(A | 3) = (0.98) (0.98) (0.98)

$P(A) = (0.20)(0.98) + (0.35)(0.98)^2 + (0.45)(0.98)^3$

= 0.1960 + 0.3361 + 0.4322 = 0.9643

Hence the probability of no defectives during a day's production is 0.9643.

6.25 A : an engineer has a bachelor's degree only

B : an engineer has a master's degree

C : an engineer is under 30 years of age

D : an engineer is over 40 years of age

(a) P(A) = 150/200 = 0.75

(b) $P(B|D) = \frac{P(B\cap D)}{P(D)} = \frac{10/200}{50/200} = 0.20$

(c) $P(C|A) = \frac{P(C\cap A)}{P(A)} = \frac{90/200}{150/200} = 0.60$

6.26 Given that

	Employed	*Unemployed*	*Total*
Males	0.40	0.10	0.50
Females	0.475	0.025	0.50
Total	0.875	0.125	1.00

Let M and F be the male and female chosen, respectively.

U = Male, female chosen is unemployed

(a) $P(M|U) = \frac{P(M\cap U)}{P(U)} = \frac{0.10}{0.125} = 0.80$

(b) $P(F|U) = \frac{P(F\cap U)}{P(U)} = 0.20$

6.27 The probability that the officer is happy and accedes to requests = 0.6 × 0.4.

The probability that the officer is unhappy and accedes to requests = 0.4 × 0.1 = 0.04.

Total probability of acceding to requests = 0.24 + 0.04 = 0.28.

The probability of his being happy if he accedes to requests = 0.24/0.28 = 0.875.

6.28 (a) P(D) = P(A) + P(B) + P(C) = 0.35 + 0.08 + 0.01 = 0.44

(b) $P(A|D) = \frac{P(A\cap D)}{P(D)} = \frac{0.35}{0.44} = 0.80$

(c) Since P(A|D) = 0.80 and P(A) = 0.35 + 0.25 = 0.80, events A and D must be independent.

6.29 Let R = Red toy is chosen and G = Green toy is chosen.

P(Both toys are R) = P(R on first choice ∩ R on second choice)

= P(R on first choice). P(R on second choice | R on first choice)

= (2/8) (1/7) = 1/28.

6.30 Let A : Executive who would remain with the company despite an equal or slightly better offer

B : Executive who has more than 10 years of service with the company

P(A and B) = P(A) P(B|A) = (120/200) (75/120) = 0.375

6.6 BAYES' THEOREM

In the 18th century, reverend Thomas Bayes, an English Presbyterian minister, raised a question: Does God really exists? To answer this question, he attempted to develop a formula to determine the probability that God does exist, based on evidence that was available to him on earth. Later, Laplace refined Bayes' work and gave it the name *Bayes' Theorem*.

Bayes' Theorem: A method to compute posterior probabilities (conditional probabilities under statistical dependence).

Bayes' theorem is useful in revising the original (prior) probability estimates of known outcomes based on additional information about these outcomes. The new estimate of original probabilities of outcomes in view of additional information is called *revised* or *posterior probabilities*.

Suppose $A_1, A_2, ..., A_n$ represent n mutually exclusive and collectively exhaustive events with marginal probabilities $P(A_1), P(A_2), ..., P(A_n)$. Let an arbitrary event, B occurred with probability, $P(B) \neq 0$. The conditional probabilities $P(B|A_1), P(B|A_2), ..., P(B|A_n)$ for event B are also known.

The revised (or posterior) probabilities, $P(A_i|B)$ of events A_i given the information that outcome B has occurred are determined by using the formula:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} \qquad (6\text{-}8)$$

Posterior Probability: A revised probability of an event obtained after getting additional information.

The formula (6-8) is called Bayes' theorem. The **posterior probabilities** $P(A_i|B)$, given that event B has occurred are the conditional probabilities of events A_i. Since events $A_1, A_2,..., A_n$ are mutually exclusive and collectively exhaustive, the event B is bound to occur with either of events $A_1, A_2, ..., A_n$. That is,

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup ... \cup (A_n \cap B)$$

As $(A_1 \cap B), (A_2 \cap B), ..., (A_n \cap B)$ are also mutually exclusive events, therefore

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + ... + P(A_n \cap B) \\ &= P(A_1)\,P(B|A_1) + P(A_2)\,P(B|A_2) + ... + P(A_n)\,P(B|A_n) \\ &= \sum_{i=1}^{n} P(A_i)\,P(B|A_i)\ a \end{aligned}$$

Using formula (6-8) for a particular event i, we have

$$\begin{aligned} P(A_i \mid B) &= \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)\,.\,P(A_i)}{P(B)} \\ &= \frac{P(B|A_i)\,.\,P(A_i)}{P(A_1)\,.\,P(B|A_1) + P(A_2)\,P(B|A_2) + \ldots + P(A_n)\,P(B|A_n)} \end{aligned}$$

Example 6.30: Suppose an item is manufactured by three machines X, Y and Z. All the three machines have equal capacity and are operated at the same rate. It is known that the percentages of defective items produced by X, Y and Z are 2, 7 and 12 per cent, respectively. All the items produced by X, Y and Z are put into one bin. From this bin, one item is drawn at random and is found to be defective. What is the probability that this item was produced on Y?

Solution: The prior probabilities of defective items produced on machines X, Y and Z are $P(X) = 1/3$; $P(Y) = 1/3$ and $P(Z) = 1/3$.

Let A be the defective item. Then, it is given that

$$P(A|X) = 0.02,\ P(A|Y) = 0.07,\ P(A|Z) = 0.12$$

Now given that the item drawn is defective, the revised probability that it was produced on machine Y is given by

$$\begin{aligned} P(Y|A) &= \frac{P(A|Y) \cdot P(Y)}{P(X)\,P(A|X) + P(Y)\,P(A|Y) + P(Z)\,P(A|Z)} \\ &= \frac{(0.07)\,.\,(1/3)}{(1/3)\,(0.02) + (1/3)\,(0.07) + (1/3)\,(0.12)} = 0.33 \end{aligned}$$

Example 6.31: Assume that a factory has two machines. Past records show that machine 1 produces 30 per cent of the items of output and machine 2 produces 70 per cent of the items. Further, 5 per cent of the items produced by machine 1 were defective and only 1 per cent produced by machine 2 were defective. If a defective item is drawn at random, what is the probability that the defective item was produced by machine 1 or machine 2?

Solution: Let A_1 and A_2 be items produced by machine 1 and 2, respectively, and D be a defective item produced either by machine 1 or machine 2.

From the data given in the problem, we know that

$$P(A_1) = 0.30,\ P(A_2) = 0.70;\ P(D \mid A_1) = 0.05,\ P(D \mid A_2) = 0.1$$

The data of the problem can now be summarized as under:

Event	*Prior Probability* $P(A_i)$ (1)	*Conditional Probability Event* $P(D\|A_i)$ (2)	*Joint Probability* $P(A_i$ and $D)$ (3)	*Posterior (Revised) Probability* $P(A_i\|D)$ $P(A_i$ and $D)$ (2) × (3)
A_1	0.30	0.05	0.015	0.015/0.022 = 0.682
A_2	0.70	0.01	0.007	0.007/0.022 = 0.318

where $P(D) = \sum_{i=1}^{2} P(D|A_i)\,P(A_i) = 0.05\times0.30 + 0.01\times0.70 = 0.22.$

Calculations in the table show that the revised probability of producing defective item by machine 2 is only 0.318 or 31.8 per cent. These calculations indicate that the defective item is more likely drawn from the output produced by machine 1.

Example 6.32: In a bolt factory, machines A, B and C manufacture 25 per cent, 35 per cent and 40 per cent of the total output respectively. Of the total of their output, 5, 4 and 2 per cent are defective bolts. A bolt is drawn at random and is found to be defective. What is the probability that it was manufactured by machines A, B or C?

[Punjab Univ., M.Com.; Madurai Univ., M.Com., 2006]

Solution: Let $A_i (i = 1, 2, 3)$ be the event of drawing a bolt produced by machine A, B and C, respectively. From the data given in the problem, we know that

$$P(A_1) = 0.25;\ P(A_2) = 0.35, \text{ and } P(A_3) = 0.40$$

Also, additional information reveals that

$$P(E|A_1) = 0.05;\ P(E|A_2) = 0.04, \text{ and } P(E|A_3) = 0.02$$

where E is the event of drawing a defective bolt

The posterior probabilities can be calculated as under:

Event	*Prior Probability* $P(A_i)$	*Conditional Probability* $P(E\|A_i)$	*Joint Probability* $P(A_i) \times P(E\|A_i)$	*Posterior Probability* $P(A_i\| E)$
A_1	0.25	0.05	0.0125	0.0125/0.0345 = 0.362
A_2	0.35	0.04	0.0140	0.0140/0.0345 = 0.406
A_3	0.40	0.02	0.0080	0.0080/0.0345 = 0.232
Total	1.00		0.0345	1.000

Since posterior probability $P(A_2 | E)$ is more among all, it is probable that defective bolt was manufactured by machine B.

Self-practice Problems 6C

6.31 A manufacturing firm produces steel pipes in three plants with daily production volumes of 500, 1000 and 2000 units respectively. According to past experience, it is known that the fractions of defective output produced by the three plants are respectively 0.005, 0.008 and 0.010. If a pipe is selected from a day's total production and found to be defective, find out (a) from which plant the pipe comes, (b) what is the probability that it came from the first plant?

[*IIT Roorkee MBA, 2004*]

6.32 In a post office, three clerks are assigned to process incoming mail. The first clerk, A, processes 40 per cent; the second clerk, B, processes 35 per cent; and the third clerk, C, processes 25 per cent of the mail. The first clerk has an error rate of 0.04, the second has an error rate of 0.06 and the third has an error rate of 0.03. A mail selected at random from a day's output is found to have an error. The postmaster wishes to know the probability that it was processed by clerk A or clerk B or clerk C.

6.33 A certain production process produces items 10 per cent of which defective. Each item is inspected before supplying to customers but 10 per cent of the time the inspector incorrectly classifies an item. Only items classified as good are supplied. If 820 items have been supplied in all, how many of them are expected to be defective?

6.34 A factory produces certain types of output by three machines. The respective daily production figures are: Machine A = 3000 units; Machine B = 2500 units; and Machine C = 4500 units. Past experience shows that 1 per cent of the output produced by machine A is defective. The corresponding fractions of defectives for the other two machines are 1.2 and 2 per cent respectively. An item is drawn at random from the day's production and is found to be defective. What is probability that it comes from the output of (a) Machine A, (b) Machine B and (c) Machine C?

6.35 In a bolt factory machines A, B and C manufacture 25 per cent, 30 per cent and 40 per cent of the total output respectively. Of the total of their output 5, 4 and 2 per cent are defective bolts. A bolt is drawn at random from the lot and is found to be defective. What are the probabilities that it was manufactured by machines A, B, or C?

6.36 In a factory manufacturing pens, machines X, Y and Z manufacture 30, 30 and 40 per cent of the total production of pens, respectively. Of their output 4, 5 and 10 per cent of the pens are defective. If one pen is selected at random, and it is found to be defective, what is the probability that it is manufactured by machine Z?

6.37 A worker-operated machine produces a defective item with probability 0.01, if the worker follows the machine's operating instruction exactly, and with probability 0.03 if he does not. If the worker follows the instructions 90 per cent of the time, what proportion of all items produced by the machine will be defective?

6.38 Medical case histories indicate that different illnesses may produce identical symptoms. Suppose a particular set of symptoms, 'H' occurs only when one of three illnesses: A, B or C occurs, with probabilities 0.01, 0.005 and 0.02 respectively. The probability of developing the symptoms H, given an illness A, B and C are 0.90, 0.95 and 0.75 respectively. Assuming that an ill person shows the symptoms H, what is the probability that a person has illness A?

Hints and Answers

6.31 Let A_1, A_2 and A_3 = production volume of plant I, II and III, respectively.

$$E = \text{defective steel pipe}$$
$$P(A_1) = 500/3500 = 0.1428;$$
$$P(A_2) = 1000/3500 = 0.2857;$$
$$P(A_3) = 2000/3500 = 0.5714$$
$$P(E \mid A_1) = 0.005,\ P(E \mid A_2) = 0.008,$$

and $$P(E \mid A_3) = 0.010.$$

$$P(A_1 \cap E) = P(A_1)\, P(E \mid A_1) = 0.1428 \times 0.005 = 0.0007;$$
$$P(A_2 \cap E) = P(A_2)\, P(E \mid A_2) = 0.2857 \times 0.008 = 0.0022$$
$$P(A_3 \cap E) = P(A_3)\, P(E \mid A_3) = 0.5714 \times 0.010 = 0.057$$
$$P(E) = P(A_1 \cap E) + P(A_2 \cap E) + P(A_3 \cap E) = 0.0007 + 0.0022 + 0.057 = 0.0599$$

(a) $$P(A_1 | E) = \frac{P(A_1 \cap E)}{P(E)} = \frac{0.0007}{0.0599} = 0.0116$$
$$P(A_2 | E) = \frac{P(A_2 \cap E)}{P(E)} = \frac{0.0022}{0.0599} = 0.0367;$$
$$P(A_3 | E) = \frac{P(A_3 \cap E)}{P(E)} = \frac{0.057}{0.0599} = 0.951$$

Since $P(A_3 \mid E)$ is highest, the defective steel pipe has most likely come from the third plant.

(b) $$P(A_1 | E) = \frac{P(A_1 \cap E)}{P(E)} = \frac{P(A_1)\, P(E|A_1)}{P(E)} = \frac{(500/3500) \times 0.005}{0.0599} = 0.0119$$

6.32 Let A, B and C = mail processed by first, second and third clerk, respectively

E = mail containing error

Given P(A) = 0.40, P(B) = 0.35 and P(C) = 0.25

P(E | A) = 0.04, P(E | B) = 0.06, and P(E | C) = 0.03

$$\therefore\ P(A \mid E) = \frac{P(A)\,P(E|A)}{P(E)}$$

$$= \frac{P(A)\,P(E|A)}{P(A)\,P(E|A) + P(B)\,P(E|B) + P(C)\,P(E|C)}$$

$$= \frac{0.40 \times 0.04}{0.40\,(0.04) + 0.35\,(0.06) + 0.25\,(0.03)} = 0.36$$

Similarly $P(B|E) = [P(B)\,P(E|B)]\,/\,P(E) = 0.47$

$P(C|E) = [P(C)\,P(E|C)]\,/\,P(E) = 0.17$

6.33 P(D) = Probability of defective item = 0.1; P(classified as good | defective) = 0.1

∴ P(G) = Probability of good item = 1 – P(D) = 1 – 0.1 = 0.9

P(classified as good | good) = 1 – P(classified as good | defective) = 1 – 0.1 = 0.9

∴ P(defective | classified as good)

$$= \frac{P(D)\cdot P(\text{classified as good}\,|\,\text{defective})}{[P(D)\cdot P(\text{classified as good}\,|\,D) + P(G)\,P(\text{classified as good}\,|\,G)]}$$

$$= \frac{0.1\times 0.1}{0.1\times 0.1 + 0.9\times 0.9} = \frac{0.01}{0.82} = 0.012.$$

6.34 (a) 0.20 (b) 0.20 (c) 0.60

6.35 P(A) = 0.37, P(B) = 0.40, P(C) = 0.23

6.36 P(Z) = 0.6639

6.37 P(A) = 0.012

6.38 P(A | H) = 0.3130

Formulae Used

1. Counting methods for determining the number of outcomes
   - Multiplication method
     (i) $n_1 \times n_2 \times \ldots \times n_k$
     (ii) $n_1 \times n_2 \times \ldots \times n_k = n^k$
     when the event in each trial is the same
   - Number of Permutations $^nP_r = \dfrac{n!}{(n-r)!}$
   - Number of Combinations $^nC_r = \dfrac{n!}{r!\,(n-r)!}$
2. Classical or *a priori* approach of computing probability of an event A

$$P(A) = \frac{\text{Number of favourable cases for A}}{\text{All possible cases}} = \frac{c(n)}{c(s)}$$

3. Relative frequency approach of computing probability of an event A in n trials of an experiment

$$P(A) = \lim_{n\to\infty} \frac{c(A)}{n}$$

4. Rule of addition of two events
   - When events A and B are mutually exclusive
     P(A or B) = P(A) + P(B)
   - When events A and B are not mutually exclusive
     P(A or B) = P(A) + P(B) – P(A and B)
5. Conditional probability
   - For statistically independent events
     $P(A|B) = P(A);\ P(B|A) = P(B)$
   - For statistically dependent events

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

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

6. Rule of multiplication of two events
   - Joint probability of independent events
     $P(A \text{ and } B) = P(A)\times P(B)$
   - Joint probability of dependent events
     $P(A \text{ and } B) = P(A|B)\times P(B)$
     $P(A \text{ and } B) = P(B|A) \times P(A)$
7. Rule of elimination
   (i) $P(B) = \Sigma P(A_i)\,P(B|A_i)$
   (ii) $P(A) = \Sigma P(B_i)\,P(A|B_i)$
8. Baye's rule $P(A_i|B) = \dfrac{P(A_i)\,P(B|A_i)}{\Sigma\,P(A_i)\,P(B|A_i)}$
9. Basic rules for assigning probabilities
   - The probability assigned to each experimental outcome
     $0 \le P(A_i) \le 1$ for all i
   - Sum of the probabilities for all the experimental outcomes
     $\Sigma\,P(A_i) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1$

   Complement of an event, $P(A) = 1 - P(\bar{A})$

Chapter Concepts Quiz

True or False

1. [T] [F] The classical approach to probability theory requires that the total number of possible outcomes be known or calculated and that each of the outcomes be equally likely.
2. [T] [F] For any two statistically independent events, P(A or B) = P(A) + P(B).
3. [T] [F] The marginal probability of an event can be formed by all the possible joint probabilities which include the event as one of the events.
4. [T] [F] The posterior probabilities help a decision-maker to update his prior probabilities by using additional experimental data.
5. [T] [F] The posterior probabilities are valid only when there are two elementary events and consistent outcomes.
6. [T] [F] When someone says that the probability of occurrence of an event is 30 per cent, he is stating a classical probability.
7. [T] [F] If $x = 4!/0!$, then x is not well defined.
8. [T] [F] Two events are mutually exclusive if their probabilities are less than one.
9. [T] [F] If events are mutually exclusive and collectively exhaustive, then posterior probabilities for these events can be equal to their prior probabilities.
10. [T] [F] *A priori* probability is estimated prior to receiving new information.
11. [T] [F] An event is one or more of the possible outcomes which result from conducting an experiment.
12. [T] [F] The condition of statistical independence arise when the occurrence of one event has no effect upon the probability of occurrence of any other event.
13. [T] [F] The collective exhaustive list of outcomes to an experiment contains every single outcome possible.
14. [T] [F] A marginal probability is also known as unconditional probability.
15. [T] [F] The relative frequency approach of assessing the probability of some event gives the greatest flexibility.

Multiple Choice Questions

16. If events are mutually exclusive, then
    (a) their probabilities are less than one
    (b) their probabilities sum to one
    (c) both events cannot occur at the same time
    (d) both of them contain every possible outcome of an experiment
17. Posterior probabilities for certain events are equal to their prior probabilities provided:
    (a) all the prior probabilities are zero
    (b) events are mutually exclusive and collectively exhaustive
    (c) events are statistically independent
    (d) none of the above
18. Two events A and B are statistically independent when
    (a) $P(A \cap B) = P(A) \times P(B)$
    (b) $P(A|B) = P(A)$
    (c) $P(A \cup B) = P(A) + P(B)$
    (d) both (a) and (b)
19. If $P(A \cap B) = A(A|B) \times P(B)$, then it implies that
    (a) both events are statistically dependent and independent
    (b) both events are statistically dependent
    (c) both events are statistically independent
    (d) none of the above
20. What is the probability that a value chosen at random from a population is larger than the median of the population?
    (a) 0.25 (b) 0.50
    (c) 0.75 (d) none of the above
21. Bayes' Theorem is useful in
    (a) revising probability estimates
    (b) computing conditional probabilities
    (c) computing sequential probabilities
    (d) none of the above
22. A probability of getting the digit 2 in a throw of unbiased dice is
    (a) zero (b) 1/2
    (c) 1/6 (d) 3/4
23. If two dice are thrown simultaneously, then the probability of getting a total of 6 will be:
    (a) 1/36 (b) 3/36
    (c) 5/36 (d) 7/36
24. A bag contains 3 red, 6 white and 7 blue balls. If two balls are drawn at random, then the probability of getting both white balls is:
    (a) 5/40 (b) 6/40
    (c) 7/40 (d) 14/40
25. What is the probability of getting an odd number in tossing a dice?
    (a) 1/6 (b) 1/3
    (c) 1/2 (d) 1
26. What is the probability of getting more than 4 in rolling a dice?
    (a) 1/6 (b) 1/3
    (c) 1/2 (d) 1

27. If the outcome is an odd number when a dice is rolled, then the probability that it is a prime number:
(a) 1/3 (b) 2/3
(c) 1/6 (d) 5/6

28. If $P(A \cap B) = 0.20$ and $P(B) = 0.80$, then $P(A \mid B)$ is
(a) 0.25 (b) 0.40
(c) 0.50 (d) 0.75

29. If $P(AB) = 0$, then the events A and B are
(a) independent (b) dependent
(c) equally likely (d) none of these

30. If $P(A \cap B) = 0.60$ and $P(A \cup B) = 0.70$ for two events A and B, then $P(A) + P(B)$ is
(a) 0.10 (b) 0.90
(c) 1.00 (d) 0.75

31. If one event is unaffected by the outcome of another event, the two events are said to be
(a) dependent (b) independent
(c) mutually exclusive (d) joint

32. If $P(A \cup B) = P(A)$, then events A and B are
(a) mutually exclusive (b) dependent
(c) independent (d) none of these

33. If probability of choosing a value at random from a particular population is larger, then the median of population is
(a) 0.25 (b) 0.50
(c) 0.75 (d) 1.00

34. The chance of rain today is 80 per cent. Which of the following best explains this statement
(a) It will rain 80 per cent today
(b) It will rain in 80 per cent of the area to which this forecast applies today
(c) In the past, weather conditions of this short have produced rain in this area 80 per cent of the time.
(d) none of these

35. Which of the following pairs of events are mutually exclusive?
(a) A contractor loses a major contract, and he increases his work force by 50 per cent.
(b) A man is older than his uncle and he is younger than his cousins.
(c) A football team loses its last game of the year, and it wins the world cup.
(d) none of these

Concepts Quiz Answers

1. T	**2.** F	**3.** T	**4.** T	**5.** F	**6.** F	**7.** F	**8.** F	**9.** T
10. T	**11.** T	**12.** T	**13.** T	**14.** T	**15.** F	**16.** (c)	**17.** (b)	**18.** (d)
19. (a)	**20.** (b)	**21.** (a)	**22.** (c)	**23.** (c)	**24.** (a)	**25.** (c)	**26.** (b)	**27.** (b)
28. (c)	**29.** (d)	**30.** (a)	**31.** (b)	**32.** (d)	**33.** (b)	**34.** (c)	**35.** (c)	

Review Self-practice Problems

6.39 Suppose a nationwide screening programme instituted through schools is being considered to uncover child abuse. It is estimated that 2 per cent of all children are subject to abuse. Further, existing screening programmers are able to determine correctly that abuse occurs 92 per cent of the time and that abuse is incorrectly suspected 5 per cent of the time.

(a) What is the probability that the results of screening indicating abuse are associated with children who are actually not abused?
(b) Based upon every 1,00,000 children screened, how many screenings can be expected to lead to a false accusation of abuse?
(c) Based upon your answer to part (a), is it valid to conclude that 73 per cent of the families not abusing children would be falsely accused? Why or why not ? *[Delhi Univ., MBA, 2004]*

6.40 If there is an increase in capital investment next year, the probability that the price of structural steel will increase is 0.90. If there is no increase in such investment, the probability of an increase is 0.40. Overall, we estimate that there is a 60 per cent chance that capital investment will increase next year.

(a) What is the overall probability of an increase in structural steel prices next year?
(b) Suppose that during the next year structural steel prices in fact increase, what is the probability that there was an increase in capital investment?

[Delhi Univ., MBA, 2000]

6.41 A product is assembled from three components X, Y and Z, and the probability of these components being defective is 0.01, 0.02 and 0.05. What is the probability that the assembled product will not be defective? *[Delhi Univ., MBA, 2001]*

6.42 A human resource manager has found it useful to categorize engineering job applicants according to their degree in engineering and relevant work experience. Out of all applicants for the job 70 per cent have a degree with or without any work experience, and 60 per cent have work experience with or without the degree. Fifty per cent of the applicants have both the degree and relevant work experience.

(a) Determine the probability that a randomly selected job applicant has either the degree or relevant work experience.
(b) What is the probability that the applicant has neither the degree nor work experience?

6.43 A salesman is found to complete a sale with 10 per cent of potential customers contacted. If the salesman randomly selects two potential customers and calls on them, then (a) what is the probability that both the calls will result in sales? and (b) what is the probability that the two calls will result in exactly one sale?

6.44 Suppose 80 per cent of the material received from a vendor is of exceptional quality, while only 50 per cent of the material received from vendor B is of exceptional quality. However, the manufacturing capacity of vendor A is limited, and for this reason only 40 per cent of the material purchased comes from vender A. The other 60 per cent comes from vendor B. An incoming shipment of material is inspected and it is found to be of exceptional quality. What is the probability that it came from vender A.

6.45 The municipal corporation routinely conducts two independent inspections of each restaurant, with the restaurant passing only if both inspectors pass it. Inspector A is very experienced, and hence, passes only 2 per cent of restaurants that actually do have rules violations. Inspector B is less experienced and passes 7 per cent restaurants with violations. What is the probability that:

(a) A reports favourable, given that B has found a violation?
(b) B reports favourable with a violation, given that inspector A passes it?
(c) A restaurant with a violation is cleared by the corporation.

6.46 If a hurricane forms in the Indian Ocean, there is a 76 per cent chance that it will strike the western coast of India. From data gathered over the past 50 years, it has been determined that the probability of a hurricane's occurring in this area in any given year is 0.85. What is the probability that a hurricane will occur in the eastern Indian Ocean and strike India this year?

6.47 A departmental store has been the target of many shoplifters during the past month, but owing to increased security precautions, 250 shoplifters have been caught. Each shoplifter's sex is noted, also noted is whether he/she was a first-time or repeat offender. The data are summarized in the table below:

Sex	*First-Time Offender*	*Repeat Offender*
Male	60	70
Female	44	76

Assuming that an apprehended shoplifter is chosen at random, find:

(a) The probability that the shoplifter is male.
(b) The probability that the shoplifter is a first-time offender, given that the shoplifter is male.
(c) The probability that the shoplifter is female, given that the shoplifter is a repeat offender.
(d) The probability that the shoplifter is female, given that the shoplifter is a first-time offender.

6.48 A doctor has decided to prescribe two new drugs to 200 heart patients in the following manner: 50 get drug A, 50 get drug B, and 100 get both. Drug A reduces the probability of a heart attack by 35 per cent drug B, reduces the probability by 20 per cent, and the two drugs, when taken together, work independently. The 200 patients were chosen so that each has an 80 per cent chance of having a heart attack. If a randomly selected patient has a heart attack, what is the probability that the patient was given both drugs?

6.49 The Deputy Commissioner of Police is trying to decide whether to schedule additional patrol units in two sensitive areas, A and B, in his district. He knows that on any given day during the past year, the probabilities of major crimes and minor crimes being committed in area A were 0.478 and 0.602, respectively, and that the corresponding probabilities in area B were 0.350 and 0.523. Assume that major and minor crimes occur independently of each other and likewise that crimes in the two areas are independent of each other.

(a) What is the probability that no crime of either type will be committed in the area A on a given day?
(b) What is the probability that a crime of either type will be committed in the area B on a given day?
(c) What is the probability that no crime of either type will be committed in either areas on a given day?

6.50 The press-room supervisor for a daily newspaper is asked to find ways to print the paper closer to distribution time, thus giving the editorial staff more leeway for last-minute changes. He has the option of running the presses at 'normal' speed or at 110 per cent of normal—'fast' speed. He estimates that these will run at the higher speed 60 per cent of the time. The roll of paper (the newsprint 'web') is twice as likely to tear at the higher speed which would mean stopping the presses temporarily

(a) If the web on a randomly-selected printing run has a probability of 0.112 of tearing, what is the probability that the web will not tear at normal speed?
(b) If the probability of tearing at fast speed is 0.20, what is the probability that a randomly-selected torn web occurred at normal speed?

[*Delhi Univ., MBA, 2001*]

6.51 The results of conducting an examination in two papers, A and B, for 20 candidates were recorded as under: 8 passed in paper A, 7 passed in paper B, 8 failed in both papers. If out of these candidates one is selected at random, find the probability that the candidate (a) passed in both A and B, (b) failed only in A, and (c) failed in A or B.

6.52 When two dice are thrown n number of times, the probability of getting at least one double six is greater than 99 per cent. What is the least numerical value of n.

6.53 It is known from past experience that a football team will play 40 per cent of its matches on artificial turf this season. It is also known that a football player's chances of incurring a knee injury are 50 per cent higher if he is playing an artificial turf instead of grass. Further, if a player's probability of knee injury on artificial turf is 0.42, what is the probability that (a) a randomly selected player incurs a knee injury, and (b) a randomly selected player with a knee injury, incurred the injury playing on grass?

6.54 In a locality of 5000 people, 1200 are above 30 years of age and 3000 are females. Out of 1200 who were above 30 years of age, 200 are females. A person is chosen at random and you are told that the person is female. What is the probability that she is above 30 years of age? *[IGNOU, 1997; Delhi Univ., MBA, 1998, 2001]*

6.55 Suppose 5 men out of 100 and 25 women out of 1000 are colour blind. A colour blind person is chosen at random. What is the probability of his being male (assuming that males and females are equal in proportion).

6.56 An organization dealing with consumer products wants to introduce a new product in the market. Based on its past experience, it has a 65 per cent chance of being successful and 35 per cent of not being successful. In order to help the organization to make a decision on the new product, that is, whether to introduce or not, it decides to get additional information on consumer attitude towards the product. For this purpose, the organization decides on a survey. In the past when a product of this type was successful, surveys yielded favourable indication 85 per cent of the time, whereas unsuccessful products received favourable survey indications 30 per cent of the time. Determine the posterior probability of the product being successful given the survey information. *[IGNOU, 1999]*

6.57 Police Head Quarter classified crime by age (in years) of the criminal and whether the crime is violent or non-violent. A total of 150 crimes were reported in the last month as shown in the table below:

Type of crime	*Age (in years)* Under 20	20–40	Over 40	*Total*
Violent	27	41	14	82
Non-violent	12	34	22	68
	39	75	36	150

(a) What is the probability of selecting a case to analyse and finding the crime was committee by someone less than 40 years old.
(b) What is the probability of selecting a case that involved a violent crime or an offender less than 20 years old?
(c) If two crimes are selected for review, then what is the probability that both are violent crimes?

6.58 With each purchase of a large pizza at a Pizza shop, the customer receives a coupon that can be scratched to see if a prize will be awarded. The odds of winning a free soft drink are 1 in 10, and the odds of winning a free large pizza are 1 in 50. You plan to eat lunch tomorrow at the shop. What is the probability.

(a) That you will win either a large pizza or a soft drink?
(b) That you will not win a prize?
(c) That you will not win a prize on three consecutive visits to the Pizza shop?
(d) That you will win at least one prize on one of your next three visits to the Pizza shop?

6.59 The boxes of men's shirts were received from the factor. Box 1 contained 25 sport shirts and 15 dress shirts. Box 2 contained 30 sport shirts and 10 dress shirts. One of the boxes was selected at random, and a shirt was chosen at random from that box to be inspected. The shirt was a sport shirt. Given this information, what is the probability that the sport shirt came from box 1?

6.60 There are four people being considered for the position of chief executive officer of enterprises. Three of the applications are over 60 years of age. Two are female, of which only one is over 60. All four applications are either over 60 years of age or female. What is the probability that a candidate is over 60 and female?

6.61 A pharmaceutical company through an advertisement in a magazine, estimates that 1 per cent of the subscribers will buy products. They also estimate that 5 per cent of non-subscribers will buy the product and that there is one chance in 20 that a person is a subscriber.

(a) Find the probability that a randomly selected person will buy the products.
(b) If a person buys the products what is the probability he subscribes to the magazine?
(c) If a person does not buy the products what is the probability he subscribes to magazine?

Hints and Answers

6.39 (a) 0.727 (b) 6740

6.40 R = rise in price of structural steel,

I = capital investment increasing.

(a) $P(R) = P(I \cap R) \cup P(\bar{I} \cap R)$

$= P(I)\,P(R|I) + P(\bar{I})\,P(R|\bar{I})$
$= 0.60 \times 0.90 + 0.40 \times 0.40 = 0.70$

(b) $P(I|R) = \dfrac{P(I \cap R)}{P(R)} = \dfrac{P(I)\,P(R|I)}{P(I)\,P(R|I) + P(\bar{I})\,P(R|\bar{I})}$

$= \dfrac{0.60 \times 0.90}{0.60 \times 0.90 + 0.40 \times 0.40}$

$= \dfrac{0.54}{0.70} = 0.77$

6.41 P (product not defective)

$= P(\bar{X})\,P(Y)\,P(Z) + P(X)\,P(\bar{Y})\,P(Z) + P(X)\,P(Y)\,P(\bar{Z})$

$= 0.99 \times 0.02 \times 0.05 + 0.01 \times 0.98 \times 0.05 + 0.01 \times 0.02 \times 0.95$

$= 0.00099 + 0.00049 + 0.00019 = 0.00167$

6.42 D = degree holders; W = with work experience

(a) $P(D \cup W) = P(D \text{ or } W) = P(D) + P(W) - P(D \cap W) = 0.70 + 0.60 - 0.50 = 0.80$

(b) $P(\bar{D} \cap \bar{W}) = 1.00 - P(D \cap W) = 1.00 - 0.80 = 0.20$

6.43 S_1, S_2 = calls resulted in sales on both the customers, respectively

(a) $P(S_1 \text{ and } S_2) = P(S_1 \cap S_2) = P(S_1)\,P(S_2) = 0.10 \times 0.10 = 0.01$

(b) $P(S_1 \cup S_2) = P(S_1 \cap\ ) \cup P(\ \cap S_2) = 0.10 \times 0.90 + 0.90 \times 0.10 = 0.18$

6.44 A = material supplied by vendor A

E = material is of exceptional quality.

$$P(A|E) = \frac{P(A \cap E)}{P(E)} = \frac{P(A)\,P(E|A)}{P(A)\,P(E|A) + P(B)\,P(E|B)} = \frac{0.40 \times 0.80}{0.40 \times 0.80 + 0.60 \times 0.50} = \frac{0.32}{0.62} = 0.516$$

6.45 (a) $P(A \mid \bar{B}) = P(A) = 0.02$

(b) $P(B \mid A) = P(B) = 0.07$

(c) $P(A \cap B) = P(A)\,P(B) = 0.02 \times 0.07 = 0.0014$

6.46 Let H = hurricane forming over Indian Ocean;

W = hurricane hits western coast of India,

$P(H \cap W) = P(H)\,P(W \mid H) = 0.76 \times 0.85 = 0.646$

6.47 M = shoplifter is male, W = shoplifter is female

F = shoplifter is first time offender,

R = shoplifter is repeat offender

(a) $P(M) = (60 + 70) \div 250 = 0.520$

(b) $P(F|M) = P(F \cap M)/P(M) = \frac{60}{250} \div \frac{130}{250} = 0.462$

(c) $P(W|R) = P(W \cap R)/P(R) = \frac{76}{250} \div \frac{146}{250} = 0.521$

(d) $P(W|F) = P(W \cap F)/P(F) = \frac{44}{250} \div \frac{104}{250} = 0.423$

6.48 H = heart attack; D = drug given

Drug	*P(D)*	*P(H | D)*
A	50/200 = 0.25	0.80 × 0.65 = 0.520
B	50/200 = 0.25	0.80 × 0.80 = 0.640
A and B	100/200 = 0.50	0.80 × 0.65 × 0.80 = 0.416

P(H ∩ D)	*P(D|H) = P(H ∩ D)/P(H)*
0.130	0.130/0.498 = 0.2610
0.160	0.160/0.498 = 0.3213
0.208	0.208/0.498 = 0.4177
P(H) = 0.498	

$P[(A \text{ and } B)/H] = P[(A \text{ and } B) \cap H]/P(H) = 0.208/0.498 = 0.417$

6.49 M_1, M_2 = major crime in district A and B, respectively

m_1, m_2 = minor crime in district A and B, respectively.

(a) $P(M_1 \cap m_1) = P(M_1) + P(m_1) - P(M_1 \cap m1)$
$= P(M_1) + P(m_1) - P(M_1)\,P(m_1)$
$= 0.478 + 0.602 - 0.478 \times 0.602$
$= 0.792$

$\therefore\ P(\bar{M}_1 \cap \bar{m}_1) = 1 - 0.792 = 0.208$

(b) $P(M_2 \cup m_2) = P(M_2) + P(m_2) - P(M_2)\,P(m_2)$
$= 0.350 + 0.523 - 0.350 \times 0.523$
$= 0.690$

(c) P (crime in A) = 0.792; P (crime in B) = 0.690

P (no crime in A and B)

= 1 – P (crime in at least A or B)

= 1 – [P (A) + P (B) – P (A and B)]

= 1 – [P (A) + P (B) – P (A) P (B)]

= 0.064

6.50 (a) Let x = P (no tear given normal speed). Then

P (tear) = P (tear | normal speed) P(normal speed) + P(tear | fast speed) P(fast speed)

$0.112 = (1 - x)(0.4) + 2(1 - x)(0.6) = 1.60 - 1.60x$

$1.6x = 1.6 - 0.112 = 1.488$ or $x = 1.488/1.6 = 0.93$

(b)

Speed	*Prob.*	*P (tour| speed)*	*P (tear and speed)*	*P (speed|tear)*
Normal	0.40	0.10	0.04	0.04/0.16 = 0.25
Fast	0.60	0.20	0.12	0.12/0.16 = 0.75
		P (tear) =	0.16	

P (normal speed | tear) = 0.25

6.51 (a) $P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{8}{20} + \frac{7}{20} - \left(1 - \frac{8}{20}\right) = \frac{3}{20}$

(b) $P(\bar{A} \cap B) = P(\bar{A}) \times P(B) = \left(\frac{12}{20}\right)\left(\frac{7}{20}\right) = 0.21$

(c) $P(\bar{A} \cup \bar{B}) = A(\bar{A}) + P(\bar{B}) - (\bar{A} \cap \bar{B}) = \frac{12}{20} + \frac{13}{20} - \frac{8}{20} = \frac{17}{20}$

6.52 Given $1 - (35/36)^n > 0.99$ or $n = 164$.

6.53 A = knee injury; B = playing on artificial turf; C = playing on grass

(a) $P(A \cap B) = P(B) \times P(A \mid B) = 0.40 \times 0.42 = 0.168$

$P(A \cap C) = P(C)\,P(A \mid C) = 0.60 \times 0.28 = 0.168$

Thus $P(A) = P(A \cap B) + P(A \cap C) = 0.168 + 0.168 = 0.336$

(b) $$P(C \mid A) = \frac{P(A \cap C)}{P(A)} = \frac{P(C) \cdot P(A|C)}{P(C)\,P(A|C) + P(B)\,P(A|B)} = \frac{0.168}{0.168 + 0.168} = \frac{1}{2}$$

6.54 P(A) = probability that a person chosen is above 30 years = 1200/5000 = 0.214

P(B) = probability that a person chosen is female = 3000/5000 = 0.60

P(A and B) = P $(A \cap B)$ = probability that a person chosen is above 30 years and a female = 200/5000 = 0.04

But $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.04}{0.06} = \frac{1}{15}$

6.55 Let M, F and C denote male, female and colour blind persons. Then

P(M|C) = 5/100 = 1/20;

P (F|C) = 25/1000 = 1/40;

P (M) = P (F) = 1/2.

Thus
$$P(C \mid M) = \frac{P(C \cap M)}{P(M)}$$
$$= \frac{P(M)\,P(M|C)}{P(M)\,P(M|C) + P(F)\,P(F|C)}$$
$$= \frac{(1/2)(1/20)}{(1/2)(1/20) + (1/2)(1/40)} = \frac{2}{3}$$

6.56 A_1, A_2 = new product is successful and failure, respectively

I = additional information

Event	*Probability* (1)	*Conditional Probability* $P(I \mid A_i)$ (2)	*Joint Probability* $P(A_i \cap I)$ (3) = (1) × (2)	*Posterior Probability* $P(A_i \mid I)$ (4) = (3)/(1)
A_1	0.65	0.85	0.552	0.84
A_2	0.35	0.30	0.105	0.16

Posterior probability of product being successful given the survey information is 0.84.

6.57 (a) P(crimes both violent and non-violent) committed by a person less than 40 years old)
$$= \frac{39}{150} + \frac{75}{100} = \frac{114}{150} = 0.76$$

(b) P(crime of violent type or offender less than 20 years old)
$$= \frac{82}{150} + \frac{39}{150} - \frac{27}{150} = \frac{94}{150} = 0.6267$$

(c) P(both crimes are of violent nature)
$$\frac{82}{150} \times \frac{81}{149} = \frac{6642}{22,300} = 0.2972$$

6.58 Let A and B represent the event of winning pizza and soft drink, respectively.

(a) $P(A \text{ or } B) = P(A)\,P(\bar{B}) + P(\bar{A})\,P(B)$
$$= \frac{1}{50} \times \frac{9}{10} + \frac{49}{50} \times \frac{1}{10} = \frac{9}{500} + \frac{49}{500} = \frac{58}{500} = 0.116$$

(b) P(no prize) $= [1 - P(A)]\,[1 - P(B)] = P(\bar{A})\,P(\bar{B})$
$$= \frac{49}{500} \times \frac{9}{10} = \frac{441}{500} = 0.882$$

(c) P(no prize on 3 visits) = $[P(\text{no prize})]^3$ = $(0.882)^2 = 0.686$

(d) P(at least are prize) = 1 – P(no prize) = 1 – 0.686 = 0.314

6.59 Given P(sport shirt) = 25/40, P(dress shirt) = 15/40 (in Box 1); P(sport short) = 30/40 P(dress shirt) = 10/40

P (shirt came from Box 1/shirt was short-shirt)
$$= \frac{P(\text{Box 1 and sport shirt})}{P(\text{sport shirt})}$$
$$= \frac{P(\text{sport shirt/Box 1})\,P(\text{Box 1})}{P(\text{Box 1})\,P(\text{sport/Box 1}) + P(\text{Box 2})\,P(\text{sport/Box 2})}$$
$$= \frac{(25/40)(1/2)}{(25/40)(1/2) + (30/40)(1/2)}$$
$$= \frac{0.625 \times 0.50}{0.625 \times 0.50 + 0.75 \times 0.50} = \frac{0.3125}{0.6875} = 0.4545$$

6.60 P(F and over 60 years) = P(F)×P(over 60 years)
$$= \frac{2}{4} \times \frac{1}{2} = 0.25$$

6.61 Given P(subscribers buy product) = 0.05;

P(Non-subscribers buy product) = 0.95

P(buy|S) = 0.01, P(not buy|S) = 0.99;

P(buy|NS) = 0.005, P(not buy|NS) = 0.995

(a) P(buy) = P(S) P(buy|S) + P(NS) P(buy|NS) = 0.05 × 0.01+ 0.95 × 0.005 = 0.00525

(b) P(S|buy) = (0.05 × 0.01)/0.00525 = 0.0952

(c) P(S|not buy) = (0.05 × 0.99)/0.00525 = 0.0498

Chapter 7

Any body can win unless there happens to be a second entry.

—George Alda

When it is not in our power to determine what is true, we ought to follow what is most probable.

— Descartes

Probability Distributions

LEARNING OBJECTIVES

After studying this chapter, you should be able to

- define the terms random variable and probability distribution.
- distinguish between discrete and continuous probability distributions.
- describe the characteristics and compute probabilities using both discrete and continuous probability distributions.
- compute expected value and variance of a random variable.
- apply the concepts of probability distributions to real life problems.

7.1 INTRODUCTION

In any probabilistic situation, choosing a strategy (course of action) may lead to a number of different possible outcomes. For example, a product whose sale is estimated around 100 units, may be equal to 100, less or more. The volume of sale (i.e., an outcome) of the product is then become an uncertain quantity and whose numerical value is termed as *random (chance* or *stochastic) variable*. In such a situation, the decision-maker may like to know the average value (payoff) of the random variable.

The list of all possible outcomes of a random variable along with their probabilities of occurrence is called *probability distribution*. The numerical value of a random variable may be different in different trials of any random experiment conducted under similar conditions. The set of all such values is called **range space** of the random variable.

Illustration If a coin is tossed twice, then the sample space, S of outcomes for this random experiment is

$$S = \{H\,H, T\,H, H\,T, T\,T\}$$

In this case, if the decision-maker is interested to know the probability distribution for the number of heads on two tosses of the coin, then a random variable (x) may be defined as

$$x = \text{number of H's occurred}$$

This random variable, x may take the values: H H = 2, H T = 1, T H = 1, T T = 0. Thus, the range space of variable values is {0, 1, 2}.

The probability of random variable, x can be found by adding the probabilities of all the simple outcomes (events) in sample space. Suppose $P(x = r)$ represents the probability of the random variable taking the value r (say, number of heads occurred, r times). Then probabilities of occurrence of head or heads in two trials are computed as:

Number of Heads (x)	*Probability of Outcome P(x)*
0	$P(x = 0) = P(TT) = P(T) \times P(T) = 0.5 \times 0.5 = 0.25$
1	$P(x = 1) = P(HT) + P(TH) = P(H) \times P(T) + P(T) \times P(H)$
	$= 0.5 \times 0.5 + 0.5 \times 0.5 = 0.25 + 0.25 = 0.50$
2	$P(x = 2) = P(HH) = P(H) \times P(H) = 0.5 \times 0.5 = 0.25$

Types of Random Variables

Discrete Random Variable: A variable that is allowed to take on only integer values.

Continuous Random Variable: A variable that is allowed to take on any value within a given range

A random variable may be either discrete or continuous. A **discrete random variable** can assume only integer values such as 0, 1, 2, . . . Quantities that are counted in whole numbers have such characteristic such as: number of letters received by a post office during a particular time period, number of machines breaking down on a given day, number of vehicles arriving at a toll bridge, number of items sold, and so on.

A **continuous random variable** can take both integer and non-integer values over a range of values. Quantities that are *measured rather than counted have this characteristic* such as time, weight, distance, temperature, amount of rainfall in a rainy season, height of individuals, and so on.

7.2 PROBABILITY DISTRIBUTION FUNCTIONS

Probability distribution functions (pdf) can be classified into two categories:

- Discrete probability distributions
- Continuous probability distributions

Discrete Probability Distribution: A probability distribution in which the random variable is permitted to take on only integer values.

In a **discrete probability distribution,** the outcomes of a random variable assume *only integer values*, such as

- A book shop has only 0, 1, 2,... copies of a particular book.
- A consumer can buy 0, 1, 2,... shirts, pants, etc.

If the random variable, x is discrete, then its probability distribution (also called *probability mass function*) must satisfy following two conditions:

(i) The probability of any specific outcome for a discrete random variable must fall between 0 and 1. Mathematically, it is stated as $0 \le P(x = r) \le 1$, for all value of r.

(ii) The sum of the probabilities of all possible values of a discrete random variable must equal 1. Mathematically, it is stated as $\sum_{\text{all } r} P(x = r) = 1$

Continuous Probability Distribution: A probability distribution in which the random variable is permitted to take any value within a given range

In a **continuous probability distribution,** the outcomes of a random variable assume both integer and non-integer values over a range of values (or interval) such as:

- Product costs and prices.
- Floor area of a house, office, etc.

If the random variable, x is continuous, then its probability density function must satisfy following two conditions:

(i) $P(x) \ge 0$; $-\infty < x < \infty$ (non-negativity condition)

(ii) $\int_{-\infty}^{\infty} P(x)\, dx = 1$ (Area under the continuous curve must total 1)

Probabilities associated with random variable values $x_1, x_2, \ldots, x_n$ in a given interval (or range), say (a, b), can be obtained by finding the area under the *pdf* between the values a and b. Mathematically, the area under *pdf* between a and b is given by

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

The expression $f(a \leq x \leq b)$ may also be expressed in terms of a differentiable distribution function. That is,

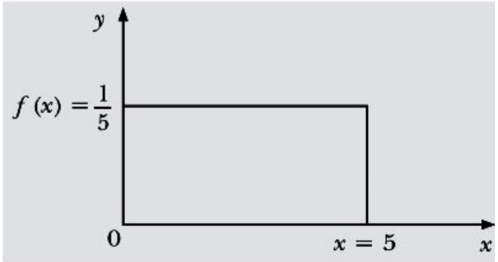
$$\frac{d}{dx}\{f(x)\} = \frac{d}{dx}\left\{\int_a^b f(x)\, dx\right\}$$

Illustration Consider the function, $f(x) = \begin{cases} a & ; 0 \leq x \leq 5 \\ 0 & ; \text{otherwise} \end{cases}$

For $f(x)$ to be a *pdf*, the condition, $\int_{-\infty}^{\infty} f(x)\, dx = 1$ must be satisfied,

which is true if $\int_0^5 a\, dx = 1$, i.e. $a = \frac{1}{5}$.

Figure 7.1
Probability Distribution Function



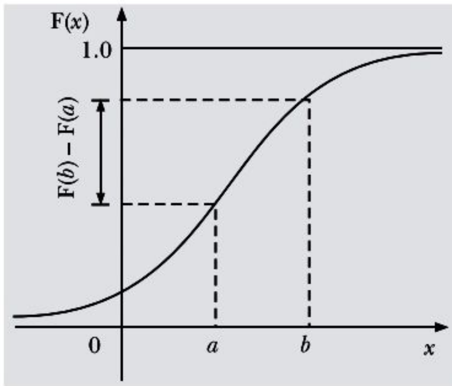
Since $a > 0$, the function, $f(x) \geq 0$. Thus, $f(x)$ satisfies both the conditions for a *pdf*. Figure 7.1 illustrates the function graphically.

7.3 CUMULATIVE PROBABILITY DISTRIBUTION FUNCTION

The cumulative probability distribution function (cdf) for the continuous random variable x $(-\infty \leq x < \infty)$ is a rule that provides the probabilities $P(x \leq r)$ for any real number r. Generally, the term cumulative probability refers to the probability that x is less than or equal to a particular value. For example, if we have three values of a random variable x as $a < b < c$, then

$$\int_a^c f(x) \geq \int_a^b f(x)$$

Figure 7.2
Cumulative Probability Distribution Function



This condition shows that *cdf* increases from left to right as shown in Fig. 7.2. Thus, the probability that the value of the random variable x is less than any real number a is given by

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

where the function $F(a)$, also called cumulative distribution (or function), represents the probability that x does not exceed a specified value 'a', and the area under the $f(x)$ curve to the left of the value a. That is, the probability of the random variable x lies at or below some specific value, a. The *cdf* has the properties

- $F(a)$ is non-decreasing function.
- $F(-\infty) = 0$ and $F(\infty) = 1$.

In general, given two real numbers a and b such that $a < b$, the probability that the value of x lies in any specified range, say between a and b, is

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

A typical *cdf* is illustrated in Fig. 7.2. However, if $f(x = a)$ and $f(x = b)$, then both of these have zero value in a continuous distribution. That is

$$f(x = a) = \int_a^a f(x) = 0$$

There is infinitely large number of possible values and the probability associated with any one of them is zero. Thus, *cdf* has the following properties:

$$\lim_{a \to \infty} F(a) = \lim_{a \to \infty} \int_{-\infty}^{a} f(x)\, dx = 1$$

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

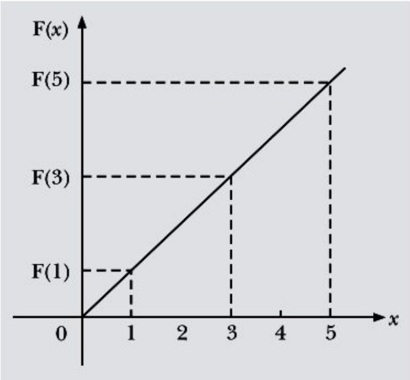


Figure 7.3 Cumulative Probability Distribution Function

Illustration For the continuous *pdf* defined as

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

the *cdf* in the range $0 \le x \le 5$ is given by

$$F(x) = \int_0^x f(x)\, dx = \int_0^x \frac{1}{5}\, dx = \frac{x}{5}$$

The *cdf* is illustrated in Fig 7.3. For example, given *pdf* the value of $f(1 \le x \le 3)$ is given by

$$f(1 \le x \le 3) = F(3) - F(1) = \frac{3}{5} - \frac{1}{5} = \frac{2}{5}$$

Illustration Let the probability distribution function of the discrete variable, x, be as follows:

Random variable :	0	1	2	3	4
Probability, $P(x = a)$:	0.10	0.20	0.40	0.20	0.10

The probability distribution function and cumulative probability distribution function of 'less than or equal to' type are shown in Table 7.1 and graphed in Fig. 7.4.

Table 7.1 Cumulative Distribution Function

Random Variable (x)	*P(x = a)*	*P(x ≤ a)*
0	0.10	0.10
1	0.20	0.30
2	0.40	0.70
3	0.20	0.90
4	0.10	1.00

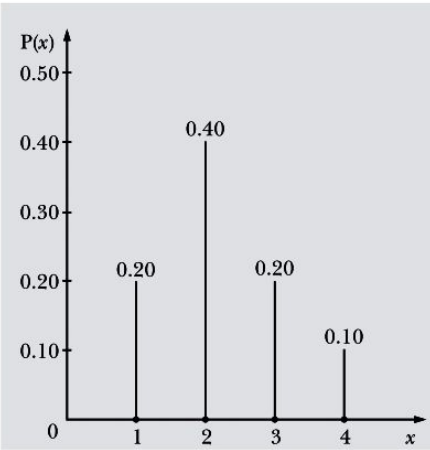


Figure 7.4 (a) Probability Distribution Function

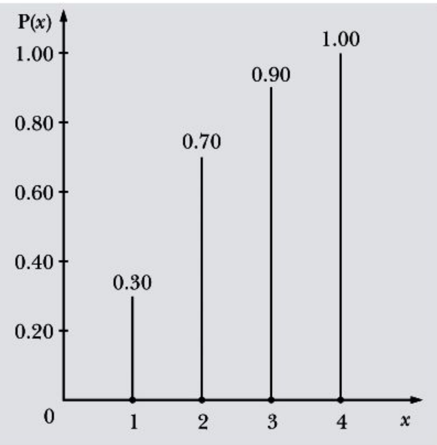


Figure 7.4 (b) Cumulative Distribution Function

From Table 7.1, the probability of x being equal to or less than a is $P(x \le a) = 0.70$. Then the probability of x being greater than a is given by $P(x > a) = 1 - P(x \le a) = 1 - 0.70 = 0.30$.

7.4 EXPECTED VALUE AND VARIANCE OF A RANDOM VARIABLE

In the same way as discussed in Chapters 3 and 4, a probability distribution is also summarized by its mean and variance.

7.4.1 Expected Value

The ***mean*** (or ***expected value***) of a random variable is the weighted average, where the possible values of random variable are weighted by the corresponding probabilities of its occurrence. If x is a random variable with possible values $x_1, x_2, ..., x_n$ occurring with probabilities $P(x_1), P(x_2),..., P(x_n)$, then the expected value of x denoted by $E(x)$ or μ is the sum of the values of the random variable weighted by the probability of its occurrence.

Expected Value of a Random Variable: A weighted average obtained by multiplying each possible value of the random variable with its probability of occurrence.

$$E(x) = \sum_{j=1}^{n} x_j P(x_j), \quad \text{provided } \sum_{j=1}^{n} P(x_j) = 1$$

Similarly, for the continuous random variable, the expected value is given by

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

where $f(x)$ is the probability distribution function.

If $E(x)$ is calculated in terms of rupees, then it is known as *expected monetary value* (EMV). For example, consider the price range of an item along with the probabilities:

Price, x :	50	60	70	80
Probability, $P(x)$:	0.2	0.5	0.2	0.1

Thus, the expected monetary value of the item is given by

$$EMV(x) = \sum_{j=1}^{n} x_j P(x_j) = 50\times0.2 + 60\times0.5 + 70\times0.2 + 80\times0.1 = ₹62.$$

7.4.2 Variance and Standard Deviation

The expected value measures the *central tendency* of a probability distribution while variance determines the *dispersion* or *variability among* possible values of random variable.

The variance, denoted by $Var(x)$ or σ^2 of a random variable, x is the squared deviation of the individual values from their expected value. That is

$$\begin{aligned} Var(x) &= E(x_j - \mu)^2 = E[(x^2 - 2x\,\mu + \mu^2)], \text{ for all } j \\ &= E(x^2) - 2\,\mu E(x) + \mu^2 = E(x^2) - \mu^2 \\ &= E(x^2) - [E(x)]^2 \end{aligned}$$

where $E(x^2) = \sum_{j=1}^{n} x_j^2 P(x_j)$ and μ or $E(x) = \sum_{j=1}^{n} x_j P(x_j)$

The variance has the disadvantage of squaring the unit of measurement. Thus, if a random variable is measured in rupees, the variance will be measured in rupee squared. This shortcoming can be avoided by using *standard deviation* (σ_x) as a measure of dispersion so as to have the same unit of measurement. That is

$$\sigma_x = \sqrt{Var(x)} = \sqrt{E(x)^2 - \mu^2}$$

7.4.3 Properties of Expected Value and Variance

The following are the important properties of an expected value of a random variable:

1. The expected value of a constant c is constant. That is, $E(c) = c$, for every constant c.
2. The expected value of the product of a constant c and a random variable x is equal to constant c times the expected value of the random variable. That is, $E(cx) = cE(x)$.
3. The expected value of a linear function of a random variable is same as the linear function of its expectation. That is, $E(a + bx) = a + bE(x)$.

4. The expected value of the product of two independent random variables is equal to the product of their individual expected values. That is, $E(xy) = E(x)\,E(y)$.
5. The expected value of the sum of the two independent random variables is equal to the sum of their individual expected values. That is, $E(x + y) = E(x) + E(y)$.
6. The variance of the product of a constant and a random variable, x is equal to the constant squared times the variance of the random variable, x. That is, $\text{Var}(cx) = c^2\text{Var}(x)$.
7. The variance of the sum (or difference) of two independent random variables equals the sum of their individual variances. That is, $\text{Var}(x \pm y) = \text{Var}(x) \pm \text{Var}(y)$.

Example 7.1: A doctor recommends a patient to take a particular diet for two weeks and there is equal chance for the patient to lose weight between 2 kg and 4 kg. What is the average amount the patient is expected to lose on this diet?

Solution: If x is the random variable, then probability density function is defined as

$$f(x) = \begin{cases} \dfrac{1}{2}, & 2 < x < 4 \\ 0, & \text{otherwise} \end{cases}$$

The amount the patient is expected to lose on the diet is

$$E(x) = \int_2^4 x.\frac{1}{2}\,dx = \left[\frac{x^2}{4}\right]_2^4 = \frac{1}{4}\left[(4)^2 - (2)^2\right] = 3 \text{ kg}$$

Example 7.2: From a bag containing 3 red balls and 2 white balls, a man is to draw two balls at random without replacement. He gains ₹20 for each red ball and ₹10 for each white one. What is the expectation of his draw.

Solution: Let x be the random variable denoting the number of red and white balls in a draw. Then x can take up the following values.

$$P(x = 2 \text{ red balls}) = \frac{{}^3C_2}{{}^5C_2} = \frac{3}{10}$$

$$P(x = 1 \text{ red and 1 white ball}) = \frac{{}^3C_1 \times {}^2C_1}{{}^5C_2} = \frac{3}{5}$$

$$P(x = 2 \text{ white balls}) = \frac{{}^2C_2}{{}^5C_2} = \frac{1}{10}$$

Thus, the probability distribution of random variable, x is

Variable :	2R	1R and 1W	2W
Gain, x :	40	30	20
Probability, $P(x)$:	3/10	3/5	1/10

Hence, expected gain is, $E(x) = 40\times(3/10) + 30\times(3/5) + 20\times(1/10) =$ ₹32.

Example 7.3: Under an employment promotion programme, it is proposed to allow sale of newspapers inside buses during off-peak hours. The vendor can purchase newspapers at a special concessional rate of ₹1.25 per copy against the selling price of ₹1.50. Any unsold copies are, however, a dead loss. A vendor has estimated the following probability distribution for the number of copies demanded.

Number of copies demanded :	15	16	17	18	19	20
Probability :	0.04	0.19	0.33	0.26	0.11	0.07

How many copies should be ordered so that his expected profit is maximum?

Solution: Profit per copy = Selling price – Purchasing price = 1.50 – 1.25 = Re 0.25. Thus,
Expected profit = Number of copies × Probability × Profit per copy

The calculations of expected profit are shown in the Table 7.2.

Table 7.2 Calculations of Expected Profit

Number of Copies Demanded (1)	*Probability* (2)	*Profit per Copy (₹)* (3)	*Expected Profit (₹)* (4) = (1)×(2)×(3)
15	0.04	0.25	15
16	0.19	0.25	76
17	0.33	0.25	**140.25**
18	0.26	0.25	117
19	0.11	0.25	52.5
20	0.07	0.25	35

The maximum profit of ₹140.25 is obtained when he stocks 17 copies of the newspaper.

Example 7.4: In a cricket match played to benefit an ex-player, 10,000 tickets are to be sold at ₹1000. The prize is ₹12,000 fridge by lottery. If a person purchases two tickets, what is his expected gain?

Solution: The gain, say x, may take one of two values: he will either lose ₹1,000 (i.e. gain will be –₹1,000) or win ₹(12,000 – 1,000) = ₹11,000, with probabilities 9,998/10,000 and 2/10,000, respectively. The probability distribution for the gain x is given below:

x	$P(x)$
–₹1000	9,998/10,000
₹11000	2/10,000

The expected gain will be

$$\begin{aligned} E(x) &= \Sigma x\, P(x) \\ &= -1000\times(9{,}998/10{,}000)+11000\times(2/10{,}000) \\ &= (-)₹997.6 \end{aligned}$$

The result implies that if the lottery were repeated an infinitely large number of times, average or expected loss will be ₹997.6.

Example 7.5: A market researcher at a major automobile company classified households by car ownership. The relative frequencies of households for each category of ownership are shown below:

Number of cars Per Household	*Relative Frequency*
0	0.10
1	0.30
2	0.40
3	0.12
4	0.06
5	0.02

Calculate the expected value and standard deviation of the random variable and interpret the result. *[Delhi Univ., MBA, 2003]*

Solution: The necessary calculations required to calculate expected and standard deviation of a random variable say, x are shown is Table 7.3.

Table 7.3 Calculations of Expected Value and Standard Deviation

Number of Cars Per Household x	*Relative Frequency, P(x)*	$x \times P(x)$	$x^2 \times P(x)$
0	0.10	0.00	0.00
1	0.30	0.30	0.30
2	0.40	0.80	1.60
3	0.12	0.36	1.08
4	0.06	0.24	0.96
5	0.02	0.10	0.50
		1.80	4.44

Expected value, $E(x) = \Sigma x\ P(x) = 1.80$. This value indicates that there are on an average 1.8 cars per household.

$$\text{Variance, } \sigma^2 = \Sigma x^2\ P(x) - [E(x)]^2 = 4.44 - (1.80)^2 = 4.44 - 3.24 = 1.20$$

$$\text{Standard deviation } \sigma = \sqrt{\sigma^2} = \sqrt{1.20} = 1.095 \text{ cars.}$$

Example 7.6: The owner of a 'Pizza Hut' has experienced that he always sells between 12 and 15 of his famous brand 'Extra Large' pizzas per day. He prepares all of them in advance and store them in the refrigerator. Since the ingredients go bad within one day, unsold pizzas are thrown out at the end of each day. The cost of preparing each pizza is ₹120 and he sells each one for ₹170. In addition to the usual cost, it cost him ₹50 per pizza that is ordered but cannot be delivered due to insufficient stock. If following is the probability distribution of the number of pizzas ordered each day, then how many 'Extra Large' pizzas should he stock each day in order to minimize expected loss.

Number of pizzas demanded	:	12	13	14	15
Probability	:	0.40	0.30	0.20	0.10

Solution: The loss matrix for the given question is shown in Table 7.4

Table 7.4 Pizza Ordered

	Pizza Ordered				
	12	*13*	*14*	*15*	
Probability → *Pizza Stocked ↓*	*0.40*	*0.30*	*0.20*	*0.10*	*Expected Loss*
12	–	100	200	300	100
13	120	–	100	200	88
14	240	120	–	100	142
15	360	240	120	–	240

Since expected loss of ₹88 is minimum corresponding to a stock level of 13 pizzas, the owner should stock 13 'Extra Large' pizzas each day.

Example 7.7: A company introduces a new product in the market and expects to make a profit of ₹2.5 lakh during the first year if the demand is 'good'; ₹1.5 lakh if the demand is 'moderate'; and a loss of ₹1 lakh if the demand is 'poor.' Market research studies indicate that the probabilities for the demand to be good and moderate are 0.2 and 0.5, respectively. Find the company's expected profit and the standard deviation.

Solution: Let x be the random variable representing profit in three types of demand. Thus, x may assume the values:

x_1 = ₹2.5 lakh when demand is good.

x_2 = ₹1.5 lakh when demand is moderate.

x_3 = ₹1 lakh when demand is poor.

Since these events (demand pattern) are mutually exclusive and exhaustive, therefore

$$P(x_1)+P(x_2)+P(x_3) = 1 \text{ or } 0.2 + 0.5 + P(x_3) = 1 \text{ or } P(x_3) = 0.3$$

Hence, the expected profit is given by

$$E(x) = 2.5 \times 0.2 + 1.5 \times 0.5 + (-1) \times 0.3 = ₹0.95 \text{ lakh}$$

Also
$$E(x^2) = x_1^2\, P(x_1) + x_2^2\, P(x_2) + x_3^2\, P(x_3)$$
$$= (2.5)^2 \times 0.2 + (1.5)^2 \times 0.5 + (-1)^2 \times 0.3 = ₹2.675 \text{ lakh}$$

Thus,
$$S.D.(x) = \sqrt{\text{Var}(x)} = \sqrt{E(x^2) - [E(x)]^2} = \sqrt{2.675 - (0.95)^2} = ₹1.331 \text{ lakh.}$$

Conceptual Questions 7A

1. Define 'random variable'. How do you distinguish between discrete and continuous random variables. Illustrate your answer with suitable examples.
2. (a) Define mathematical expectation of a random variable.
   (b) Explain what do you mean by the term 'mathematical expectation'. How is it useful for a businessman? Give an example to illustrate its usefulness. *[Delhi Univ., MBA, 2007]*
3. What is meant by probability distribution of a random variable? Distinguish between probability density function and probability mass function. Illustrate with examples.
4. What do you understand by the expected value of a random variable?
5. What are the properties of expected value and variance of a random variable?

Self-practice Problems 7A

7.1 Anil company estimates the net profit on a new product it is launching to be ₹30,00,000 during the first year if it is 'successful'; ₹10,00,000 if it is 'moderately successful'; and a loss of ₹10,00,000 if it is 'unsuccessful'. The firm assigns the following probabilities to its first year prospects for the product: Successful: 0.15 and moderately successful: 0.25. What are the expected value and standard deviation of first year net profit for this product? *[Delhi Univ., MBA, 2005]*

7.2 If the probability that the value of a certain stock will remain the same is 0.46, the probability that its value will increase by ₹0.50 or ₹ 1.00 per share are respectively 0.17 and 0.23, and the probability that its value will decrease ₹0.25 per share is 0.14, what is the expected gain per share?

7.3 A box contains 12 items of which 3 are defective. A sample of 3 items is selected at random from this box. If x represents the number of defective items of 3 selected items, describe the random variable x completely and obtain its expectation.

7.4 Fifty per cent of all automobile accidents lead to property damage of ₹100, forty per cent lead to damage of ₹500, and ten per cent lead to total loss, that is, damage of ₹1800. If a car has a 5 per cent chance of being in an accident in a year, what is the expected value of the property damage due to that possible accident?

7.5 The probability that there is at least one error in an account statement prepared by A is 0.2 and for B and C it is 0.25 and 0.4 respectively. A, B and C prepared 10, 16 and 20 statements, respectively. Find the expected number of correct statements in all.

7.6 A lottery sells 10,000 tickets at ₹1 per ticket, and the prize of ₹5000 will be given to the winner of the first draw. Suppose you have bought a ticket, how much should you expect to win?

7.7 The monthly demand for transistors is known to have the following probability distribution.

Demand (n) :	1	2	3	4	5	6
Probability (P) :	0.10	0.15	0.20	0.25	0.18	0.12

Determine the expected demand for transistors. Also obtain the variance. Suppose the cost (C) of producing 'n' transistors is given by the relationship, C = 10,000 + 500n. Determine the expected cost.

7.8 A bakery has the following schedule of daily demand for cakes. Find the expected number of cakes demanded per day.

Number of Cakes Demanded	*Probability*
0	0.02
1	0.07
2	0.09
3	0.12
4	0.20
5	0.20
6	0.18
7	0.10
8	0.01
9	0.01

7.9 A consignment of machine parts is offered to two firms, A and B, for ₹75,000. The following table shows the probabilities at which firms A and B will be able to sell the consignment at different prices.

Probability	*Price (in ₹) at which the Consignment Can be Sold*			
	60,000	70,000	80,000	90,000
A	0.40	0.30	0.20	0.10
B	0.10	0.20	0.50	0.20

Which firm, A or B, will be more inclined towards this offer?

7.10 An industrial salesman wants to know the average number of units he sells per sales call. He checks his past sales records and comes up with the following probabilities:

Sales (units) :	0	1	2	3	4	5
Probability :	0.15	0.20	0.10	0.05	0.30	0.20

You are expected to help the salesman in his objective.

7.11 A survey conducted over the last 25 years indicated that in 10 years the winter was mild, in 8 years it was cold, and in the remaining 7 it was very cold. A company sells 1000 woolen coats in a mild year, 1300 in a cold year, and 2000 in a very cold year. You are required to find the yearly expected profit of the company if a woollen coat costs ₹173 and it is sold to stores for ₹248.

Hints and Answers

7.1

x :	3	1	– 1
P(x) :	0.15	0.25	1 – 0.15 – 0.25 = 0.60

E(x) = ₹0.10 million, Var (x) = ₹2.19 million, and σx = ₹1.48 million.

7.2 ₹ 0.28

7.3

x :	0	1	2	3
P(x) :	27/64	27/64	9/64	1/64;

E(x) = 0.75

7.4

x :	100	500	1,800
P(x) :	0.50	0.40	0.10

E(x) = ₹430; 0.5 E(x) = ₹215

7.5 P(A)= 0.2; P(B) = 0.25; P(C) = 0.4; and $P(\bar{A}) = 0.8$; $P(\bar{B}) = 0.75$; $P(\bar{C}) = 0.6$

$$E(x) = x_1 P(\bar{A}) + x_2 P(\bar{B}) + x_3 P(\bar{C}) = 32$$

7.6 $P(\text{Win}) = \dfrac{9999}{10{,}000}$ and $\dfrac{1}{1000}$

$$E(x) = -1 \times \frac{9999}{10{,}000} + 4999 \times \frac{1}{1000} = ₹3.9991$$

7.7 Expected demand for transistors, E(n) = $\Sigma\, np$ = 3.62

E(C) = (10,000 + 500n) = 10,000 + 50 E(n) = ₹11,810.

7.8 E(x) = 508.

7.9 EMV (A) = 6 × 0.4 + 7 × 0.3 + 8 × 0.2 + 9 × 0.1 = ₹70,000.

EMV (B) = 6 × 0.1 + 7 × 0.2 + 8 × 0.5 + 9 × 0.2 = ₹78,000.

Firm B will be more inclined towards the offer.

7.10 E (x) = 2.75

7.11

State of Nature	*Mild*	*Cold*	*Very Cold*
Prob. P(x)	0.40	0.32	0.28
Sale of coat	1000	1300	2000
Profit, x	1000× (248 – 173)	1300× (248 – 173)	2000× (248 – 173)

E(Profit) = ₹1,03,200

7.5 DISCRETE PROBABILITY DISTRIBUTIONS

7.5.1 Binomial Probability Distribution

Binomial probability distribution is a widely used probability distribution for a discrete random variable. This distribution describes data resulting from an experiment called a *Bernoulli process* (named after Jacob Bernoulli, 1654–1705, a Swiss mathematician). For each trial of an experiment, there are *only two possible mutually exclusive outcomes* such as defective or good, head or tail, zero or one, boy or girl. In such cases, the outcome of interest is referred to as a '*success*' and the other as a '*failure*'.

Bernoulli process: *It is a process wherein an experiment is performed repeatedly, yielding either a success or a failure in each trial and where there is absolutely no pattern in the occurrence of successes and failures. That is, the occurrence of a success or a failure in a particular trial does not affect, and is not affected by the outcomes in any previous or subsequent trials. The trials are independent.*

Conditions for Binomial Experiment

The Bernoulli process involving a series of independent trials is based on following conditions:

(i) There are only two mutually exclusive and collective exhaustive outcomes of the random variable and one of them is referred to as a *success* and the other as a *failure*.

(ii) The random experiment is performed under the same conditions for a fixed and finite (also discrete) number of times, say n. Each observation of the random variable in a random experiment is called a *trial*. Each trial generates either a *success* denoted by p or *a failure* denoted by q.

(iii) The outcome (i.e., success or failure) of any trial is not affected by the outcome of any other trial. That is, outcomes are assumed to be independent of each other.

(iv) The probability of both success (p), and failure ($q = 1 - p$) remains constant from trial to trial.

To understand the Bernoulli process, consider the coin tossing problem where 3 coins are tossed simultaneously or one after other. Suppose we are interested to know the probability of two heads. The possible sequence of outcomes involving two heads can be obtained in the following three ways: HHT, HTH, THH.

The probability of outcome in each of these sequences can be found by using the multiplication rule for independent events. Let the probability of a head be p and the probability of tail be q. Then probability of each sequence can be written as

$$ppq,\ pqp\ \ qpp.$$

Since probability of H or T is same, therefore each of these probabilities can be written as p^2q.

Bernoulli Process: A process in which each trial has only two possible outcomes, the probability of the outcome at any trial remains fixed over time, and the trials are statistically independent.

Since three sequences correspond to the same event '2 heads', therefore the probability of 2 heads in 3 tosses is obtained by using the addition rule of probabilities for mutually exclusive events. Since the probability of each sequence is same, we can multiply p^2q (probability of one sequence) by 3 (number of possible sequences or orderings of 2 heads). Hence,

$$P(x = 2 \text{ heads}) = 3p^2q = {}^3C_2\, p^2q$$

It may be noted that the binomial coefficient ${}^3C_2 = 3$ represents the number of ways that three symbols, of which two are alike (i.e., 2H and one T), can be ordered (or arranged). In general, the binomial coefficient ${}^nC_r$ represents the number of ways that n symbols, of which r are alike, can be ordered.

Since outcomes H and T are equally likely and mutually exclusive, therefore $p = 0.5$ and $q = 0.5$ for a toss of the coin. Thus, the probability of 2 heads in 3 tosses is

$$P(x = 2 \text{ heads}) = {}^3C_2(0.5)^2\,(0.5) = 3(0.25)\,(0.5) = 0.375$$

Binomial Probability Function

In general, for a binomial random variable, x the probability of success (occurrence of desired outcome) r number of times in n independent trials, regardless of their order of occurrence is given by

$$P\,(x = r \text{ successes}) = {}^nC_r\, p^r q^{n-r}$$

$$= \frac{n!}{r!\,(n-r)!} p^r q^{n-r},\ r = 0, 1, 2, ..., n \qquad (7\text{-}1)$$

where n is number of trials or sample size, p is probability of success, $q = (1 - p)$ is probability of failure, x is the discrete binomial random variable and r is the number of successes in n trials.

In formula (7-1), the term $p^r q^{n-r}$ represents the probability of one sequence of outcomes, where r number of outcomes (called successes) occur in n trials in a particular sequence, while the term ${}^nC_r$ represents the number of possible sequences (combinations) of r successes that are possible out of n trials.

Characteristics of the Binomial Distribution

The expression (7-1) is known as **binomial distribution** with parameters n and p. Different values of n and p identify different binomial distributions leading to different probabilities of r-values. The *mean* and *standard deviation* of a binomial distribution are computed as follows:

Mean, $\mu = np$,

Standard deviation, $\sigma = \sqrt{n\,p\,q}$

Knowing the values of first two central moments $\mu_0 = 1$ and $\mu_1 = 1$, other central moments are given by

$$\text{Second moment, } \mu_2 = npq$$
$$\text{Third moment, } \mu_3 = npq(q - p)$$
$$\text{Fourth moment, } \mu_4 = 3n^2p^2q^2 + npq(1 - 6pq)$$

so that
$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{q-p}{\sqrt{npq}}, \text{ where } \beta_1 = \frac{n^2p^2q^2(q-p)^2}{n^3p^3q^3}$$

Binomial Distribution: A discrete probability distribution of outcomes of an experiment known as a Bernoulli process.

and
$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{1-6pq}{npq}, \text{ where } \beta_2 = \frac{3n^2p^2q^2 + npq(1-6pq)}{n^2p^2q^2}$$

For a binomial distribution, *variance<mean*. This distribution is unimodal when np is an integer number, and mean = mode = np.

A binomial distribution satisfies both the conditions of *pdf* because

$$P(x = r) \geq 0, \quad \text{for all} \quad r = 0, 1, 2, \ldots, n$$

$$\sum_{r=0}^{n} P(x = r) = \sum_{r=0}^{n} \left[{}^nC_r\, p^r\, q^{n-r}\right] = (p + q)^n = 1$$

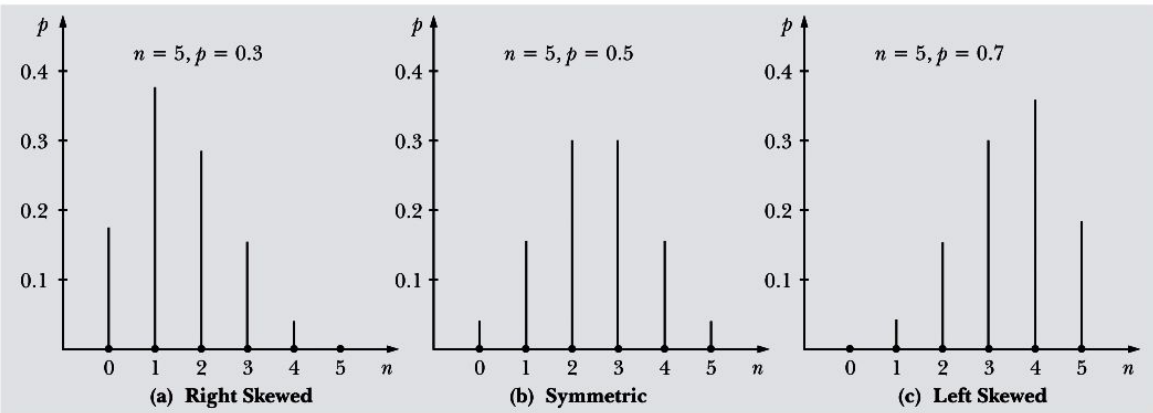
Plotting the Binomial Distributions

When value of parameters n and p in the binomial distribution are subject to a change, the characteristics of resulting binomial distributions can be shown graphically as follows:

(i) Figure 7.5 illustrates the general shape of a family of binomial distributions with constant $n = 5$ and p varies from 0.3 to 0.7.

In all three cases shown in Fig. 7.5, the skewness varies with the value of p. When p is small (i.e. $p < 0.5$), the distribution is skewed to the right. When p and q are equal (i.e. $p = q = 0.5$), the distribution is symmetric. When p is large (i.e. $p > 0.5$), the distribution is skewed to left.

Figure 7.5
Bionomial Distributions with Constants n and Variable p



The probability of success is always close to the mean value and always increases as p increases. The variance is largest when p and q are equal. The smaller the variance, larger is the probability that the value of random variable, x falls close to the mean value.

(ii) If p remains constant but n is increased, then for any value of p other than 0.5, the binomial distribution approaches symmetry. In general, smaller the value of p, the larger sample size is necessary for the symmetry of distribution.

Fitting a Binomial Distribution

A binomial distribution can be fitted to the observed values in the data set as follows:

- Find the value of p and q. If one of these is known, the other can be obtained by using the relationship $p + q = 1$.
- Expand $(p + q)^n = p^n + {}^nC_1\, p^{n-1} q + {}^nC_2\, p^{n-2} q^2 + ... + {}^nC_r\, p^{n-r} q^r + ... + {}^nC_n\, q^n$ using the concept of binomial theorem.
- Multiply each term in the expansion by the total number of frequencies, N to obtain the expected frequency for each of the random variable value.

The following recurrence relation can be used for fitting of a binomial distribution:

$$P(x = r) = {}^nC_r\, p^r q^{n-r}$$

$$P(x = r + 1) = {}^nC_{r+1}\, p^{r+1} q^{n-r-1}$$

$$\therefore \quad \frac{f(r+1)}{f(r)} = \frac{p}{q}\frac{n-r}{r+1} \text{ or } P(r + 1) = \frac{p}{q}\frac{n-r}{r+1} f(r)$$

For $r = 0$, $\quad P(x = 1) = \frac{p}{q} n f(0)$

For $r = 1$, $\quad P(x = 2) = \frac{p}{q}\frac{n-1}{2} P(x = 1) = \left(\frac{p}{q}\right)^2 \frac{n(n-1)}{2!} P(x = 0)$

For $r = 2$, $\quad P(x = 3) = \frac{p}{q}\frac{n-2}{3} P(x = 2) = \left(\frac{p}{q}\right)^3 \frac{n(n-1)(n-2)}{3!} P(x = 0) \quad (7\text{-}2)$

and so on.

In formula (7-2), we need to calculate $P(x = 0)$, which is equal to q^n, where q can be calculated from the given data.

Example 7.8: If on an average 8 ships out of 10 arrive safely at a port, find the mean and standard deviation of the number of ships arriving safely out of a total of 1600 ships.

[*Kurukshetra Univ., MBA, 2004*]

Solution: Probability of safe arrival, $p = 8/10 = 0.8$ and $q = 1 - p = 1 - 0.8 = 0.2$.

Mean number of ships returning safely: $m(= np) = 1600 \times 0.8 = 1280$.

Standard deviation, $\sigma = \sqrt{npq} = \sqrt{1600 \times 0.8 \times 0.2} = 16$.

Example 7.9: A brokerage survey reports that 30 per cent of individual investors have used a discount broker, i.e. one which does not charge the full commission. In a random sample of 9 individuals, what is the probability that

(a) exactly two of the sampled individuals have used a discount broker?
(b) not more than three have used a discount broker
(c) at least three of them have used a discount broker

Solution: The probability that individual investors have used a discount broker is $p = 0.30$, and therefore $q = 1 - p = 0.70$

(a) Probability that exactly 2 of the 9 individual have used a discount broker is given by

$$P(x = 2)\, {}^9C_2 = (0.30)^2 (0.70)^7 = \frac{9!}{(9-2)!\,2!} (0.30)^2 (0.70)^7$$

$$= \frac{9 \times 8}{2} \times 0.09 \times 0.082 = 0.2656$$

(b) Probability that out of 9 randomly selected individuals not more than three have used a discount broker is given by

$$P(x \le 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$= (0.30)^0 (0.70)^9 + (0.30)(0.70)^8 + (0.30)^2 (0.70)^7 + (0.30)^3 (0.70)^6$$

$$= 0.040 + 9 \times 0.30 \times 0.058 + 36 \times 0.09 \times 0.082 + 84 \times 0.027 \times 0.118$$

$$= 0.040 + 0.157 + 0.266 + 0.268 = 0.731$$

(c) Probability that out of 9 randomly selected individuals, at least three have used a discount broker is given by

$$P(x \geq 3) = 1 - P(x < 3) = 1 - [P(x = 0) + P(x = 1) + P(x = 2)]$$
$$= 1 - [0.040 + 0.157 + 0.266] = 0.537$$

Example 7.10: Mr Gupta applies for a personal loan of ₹1,50,000 from a nationalized bank to repair his house. The loan offer informed him that over the years, bank has received about 2920 loan applications per year and that the probability of approval was, on average, above 0.85.

(a) Mr Gupta wants to know the average and standard deviation of the number of loans approved per year.

(b) Suppose bank actually received 2654 loan applications per year with an approval probability of 0.82. What are the mean and standard deviation now?

Solution: (a) Assuming that approvals are independent from loan to loan, and that all loans have the same 0.85 probability of approval. Then

Mean, $\mu = np = 2920 \times 0.85 = 2482$

Standard deviation, $\sigma = \sqrt{npq} = \sqrt{2920 \times 0.85 \times 0.15} = 19.295$

(b) Mean, $\mu = np = 2654 \times 0.82 = 2176.28$

Standard deviation, $\sigma = \sqrt{npq} = \sqrt{2654 \times 0.82 \times 0.18} = 19.792$

Example 7.11: Suppose 10 per cent of new scooters will require warranty service within the first month of its sale. A scooter manufacturing company sells 1000 scooters in a month,

(a) Find the mean and standard deviation of scooters that require warranty service.

(b) Calculate the moment coefficient of skewness and kurtosis of the distribution.

Solution: Given that $p = 0.10$, $q = 1 - p = 0.90$ and $n = 1000$

(a) Mean, $\mu = np = 1000 \times 0.10 = 100$ scooters

Standard deviation, $\sigma = \sqrt{npq} = \sqrt{1000 \times 0.10 \times 0.90} = 10$ scootors (approx.)

(b) Moment coefficient of skewness

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{0.90 - 0.10}{9.48} = 0.084$$

Since γ_1 is more than zero, the distribution is positively skewed.

Moment coefficient of kurtosis, $\gamma_2 = \beta_2 - 3$

$$= \frac{1-6pq}{npq} = \frac{1-6(0.10)(0.90)}{90} = \frac{0.46}{90} = 0.0051$$

Since γ_2 is positive, the distribution is platykurtic.

Example 7.12: The incidence of occupational disease in an industry is such that the workers have 20 per cent chance of suffering from it. What is the probability that out of six workers 4 or more will come in contact of the disease?

[Lucknow Univ., MBA, 2006; Delhi Univ., MBA, 2002]

Solution: The probability of a worker suffering from the disease is, $p = 20/100 = 1/5$. Therefore $q = 1 - p = 1 - (1/5) = 4/5$.

The probability of 4 or more, that is, 4, 5 or 6 coming in contact of the disease is given by

$$P(x \geq 4) = P(x = 4) + P(x = 5) + P(x = 6)$$
$$= {}^6C_4\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)^2 + {}^6C_5\left(\frac{1}{5}\right)^5\left(\frac{4}{5}\right) + {}^6C_4\left(\frac{1}{5}\right)^6$$
$$= \frac{15 \times 16}{15625} + \frac{6 \times 4}{15625} + \frac{1}{15625} = \frac{1}{15625}(240 + 24 + 1)$$
$$= \frac{265}{15625} = 0.01695$$

Hence the probability that out of 6 workers 4 or more will come in contact of the disease is 0.01695.

Example 7.13: A multiple-choice test contains 8 questions with 3 answers to each question (of which only one is correct). A student answers each question by rolling a balanced dice and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4, and the third answer if he gets 5 or 6. To get a distinction, the student must secure at least 75 per cent correct answers. If there is no negative marking, what is the probability that the student secures a distinction?

Solution: Probability of a correct answer, p is one in three so that $p=1/3$ and probability of wrong answer $q = 2/3$.

The required probability of securing a distinction (i.e., of getting the correct answer of at least 6 of the 8 questions) is given by

$$P(x \geq 6) = P(x = 6) + P(x = 7) + P(x = 8)$$

$$= {}^8C_6\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^2 + {}^8C_7\left(\frac{1}{3}\right)^7\left(\frac{2}{3}\right) + {}^8C_8\left(\frac{1}{3}\right)^8$$

$$= \left(\frac{1}{3}\right)^6\left[{}^8C_6\left(\frac{2}{3}\right)^2 + {}^8C_7\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + {}^8C_8\left(\frac{1}{3}\right)^2\right]$$

$$= \frac{1}{729}\left[28 \times \frac{4}{9} + 8 \times \frac{2}{9} + \frac{1}{9}\right] = \frac{1}{729}(12.45 + 0.178 + 0.12) = 0.0196$$

Example 7.14: The screws produced by a certain machine were checked by examining the number of defectives in a sample of 12. The following table shows the distribution of 128 samples according to the number of defective items they contained:

No. of defectives in a sample of 12 :	0	1	2	3	4	5	6	7
No. of samples :	7	6	19	35	30	23	7	1 = 128

(a) Fit a binomial distribution and find the expected frequencies if the chance of machine being defective is 0.5.
(b) Find the mean and standard deviation of the fitted distribution. [*Delhi Univ., MBA, 2003*]

Solution: (a) The probability of a defective screw is, $p = 1/2$ and therefore $q = 1 - p = 1/2$; N = 128. Since there are 8 terms, therefore $n = 7$. Thus, the probability that the defective items are 0, 1, 2,..., 7 is given by

$$(p + q)^n = p^n + {}^nC_1 p^{n-1} q + {}^nC_2 p^{n-2} q^2 + \ldots + {}^nC_7 p^{n-7} q^7$$

or

$$\left(\frac{1}{2}+\frac{1}{2}\right)^7 = \left(\frac{1}{2}\right)^7 + {}^7C_1\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right) + {}^7C_2\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^2 + \ldots + {}^7C_7\left(\frac{1}{2}\right)^7$$

$$= \left(\frac{1}{2}\right)^7 [1 + 7 + 21 + 35 + 35 + 21 + 7 + 1]$$

For obtaining the expected frequencies, multiply each term by N = 128. That is,

$$128\left(\frac{1}{2}+\frac{1}{2}\right)^7 = 128 \times \frac{1}{128}(1 + 7 + 21 + 35 + 35 + 21 + 7 + 1)$$

Thus, the expected frequencies are

x :	0	1	2	3	4	5	6	7
f :	1	7	21	35	35	21	7	1

(b) The mean of binomial distribution is given by np and standard deviation by $\sqrt{npq}$. Given that, $n = 7, p = q = 1/2$. Thus

$$\text{Mean} = np = 7 \times (1/2) = 3.5$$

$$\text{Standard deviation} = \sqrt{npq} = \sqrt{7 \times (1/2) \times (1/2)} = \sqrt{1.75} = 1.32.$$

Conceptual Questions 7B

6. (a) Define binomial distribution stating its parameters, mean, and standard deviation, and give two examples where such a distribution is ideally suited.
(b) Define binomial distribution. Point out its chief characteristics and uses. Under what conditions does it tend to Poisson distribution?

7. For a binomial distribution, is it true that the mean is the most likely value? Explain.

8. Demonstrate that the binomial coefficient ${}^nC_r$ equals ${}^nC_{n-r}$ and illustrate this with a specific numerical example.

9. What assumptions must be met for a binomial distribution to be applied to a real life situation?

10. What is meant by the term parameter of a probability distribution? Relate the concept to the binomial distribution?

11. What information is provided by the mean, standard deviation, and central moments of the binomial distribution?

12. What is a binomial coefficient and illustrate this with a specific numerical example.

Self-practice Problems 7B

7.12 The normal rate of infection of a certain disease in animals is known to be 25 per cent. In an experiment with 6 animals injected with a new vaccine it was observed that none of the animals caught the infection. Calculate the probability of the observed result.

7.13 Out of 320 families with 5 children each, what percentage would be expected to have (i) 2 boys and 3 girls, (ii) at least one boy? Assume equal probability for boys and girls.

7.14 The incidence of a certain disease is such that on an average 20 per cent of workers suffer from it. If 10 workers are selected at random, find the probability that (*i*) exactly 2 workers suffer from the disease, (*ii*) not more than 2 workers suffer from the disease. Calculate the probability up to fourth decimal place.
[*MD Univ., M.Com., 2008*]

7.15 The mean of a binomial distribution is 40 and standard deviation 6. Calculate n, p, and q.
[*Delhi Univ., MBA, 2008*]

7.16 A student obtained answers with mean $\mu = 2.4$ and variance $\sigma^2 = 3.2$ for a certain problem given to him using binomial distribution. Comment on the result.

7.17 The probability that an evening college 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. [*Madras Univ., M.Com., 2007*]

7.18 The normal rate of infection of a certain disease in animals is known to be 25 per cent. In an experiment with 6 animals injected with a new vaccine it was observed that none of the animals caught infection. Calculate the probability of the observed result.

7.19 Is there any inconsistency in the statement that the mean of a binomial distribution is 20 and its standard deviation is 4? If no inconsistency is found what shall be the values of p, q and n?

7.20 Find the probability that in a family of 5 children there will be (i) at least one boy, (ii) at least one boy and one girl (Assume that the probability of a female birth is 0.5).

7.21 A famous advertising slogan claims that 4 out of 5 housewives cannot distinguish between two particular brands of butter. If this claim is valid and 5000 housewives are tested in groups of 5, how many of these groups will contain 0, 1, 2, 3, 4, and 5 housewives who do not distinguish between the two products? Assume that the capacity to distinguish between the two brands is randomly distributed so that Bernoulli trial conditions are satisfied.

7.22 A supposed coffee connoisseur claims that he can distinguish between a cup of instant coffee and a cup of percolator coffee 75 per cent of the time. It is agreed that his claim will be accepted if he correctly identifies at least 5 out of 6 cups. Find (a) his chance of having the claim accepted if he is in fact only guessing, and (b) his chance of having the claim rejected when he does have the ability he claims.

Hints and Answers

7.12 Let P denote infection of the disease. Then $p = 25/100 = 1/4$ and $q = 3/4$.

$$P(x = 0) = {}^6C_0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^6 = \frac{729}{4096}$$

7.13 (i) Given $p = q = 1/2$

$$P(\text{boy} = 2) = {}^5C_2 p^2 q^3 = {}^5C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 = \frac{5}{16} \text{ or } 31.25 \text{ per cent}$$

(ii) $P(\text{boy} \geq 1) = 1 - {}^5C_0 q^5 = 1 - \frac{1}{32} = \frac{31}{32}$ or 97 per cent.

7.14 Probability that a worker suffers from a disease, $P = 1/5$ and $q = 4/5$.

$$P(x = r) = {}^nC_r q^{n-r} p_r = {}^{10}C_r\left(\frac{1}{5}\right)^r\left(\frac{4}{5}\right)^{10-r}$$

$$= {}^{10}C_r \frac{4^{10-r}}{5^{10}};\ r = 0, 1, 2, \ldots, 10$$

(i) $P(x = 2) = {}^{10}C_2 \frac{4^{10-2}}{5^{10}} = 0.302$

(ii) $P(x = 0) + P(x = 1) + P(x = 2)$

$$= \frac{1}{5^{10}}({}^{10}C_0\, 4^{10} + {}^{10}C_1\, 4^9 + {}^{10}C_2\, 4^8) = 0.678.$$

7.15 Given $\mu = np = 40$ and $\sigma = \sqrt{npq} = 6$. Squaring σ, we get $npq = 36$ or $40q = 36$ or $q = 0.9$. Then $p = 1 - q = 0.28$.
Since $np = 40$ or $n = 40/p = 40/0.1 = 400$.

7.16 Given $\sigma^2 = npq = 3.2$ and $\mu = np = 2.4$. Then $2.4q = 3.2$ or $q = 3.2/2.4 = 1.33$ (inconsistent result)

7.17 Given $p = 0.4$ and $q = 0.6$

(a) $P(x = \text{no graduate}) = {}^5C_0(0.4)(0.6)^5 = 1 \times 1 \times 0.0777 = 0.0777$

(b) $P(x = 1) = {}^5C_1(0.4)^1(0.6)^4 = 0.2592$

(c) $P(x \geq 1) = 1 - P(x = 0) = 1 - 0.0777 = 0.9223$

7.18 Probability of infection of disease $= 25/100 = 0.25$; $q = 1 - p = 0.75$.

The first term in the expansion of $(q + p)^n = \left(\frac{3}{4} + \frac{1}{4}\right)^6$ is ${}^6C_0\left(\frac{3}{4}\right)^6 = 0.177$, which is also the required probability.

7.19 Given $\mu = np = 20$ and $\sigma = \sqrt{npq} = 4$ or $npq = 16$ or $20q = 16$ or $q = 16/20 = 0.80$ and then $p = 1 - q = 0.20$. Hence $npq = 16$ gives $n = 16/pq = 16/(0.20 \times 0.80) = 100$.

7.20 Since $p = q = 0.5$, therefore

(i) $P(\text{boy} = 0) = {}^5C_0(0.5)^0(0.5)^5 = 0.031$
P(at least one boy) $= 1 - 0.031 = 0.969$

(ii) P(at least 1B and 1G) $= {}^5C_1(0.5)^1(0.5)^4 + {}^5C_2(0.5)^2(0.5)^3 + {}^5C_3(0.5)^3(0.5)^2 + {}^5C_4(0.5)^4(0.5)$

$$= \frac{5}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} = \frac{30}{32}$$

7.21 p = probability that 4 out of 5 cannot distinguish between two brands $= 4/5 = 0.8$ so that $q = 1 - p = 1/5 = 0.2$
Expected distribution containing 0, 1, ..., 5 housewives who do not distinguish between two brands
$= {}^5C_0(0.2)^5 + {}^5C_1(0.8)(0.2)^4 + {}^5C_2(0.8)^2(0.2)^3 + {}^5C_3(0.8)^3(0.2)^2 + {}^5C_4(0.8)^4(0.2)^1 + {}^5C_5(0.8)^5$
Thus required number in each group would be
$5000(0.2)^5$; $5000\ {}^5C_1(0.8)(0.2)^4$;
$5000\ {}^5C_2(0.8)^2(0.2)^3$; $5000\ {}^5C_3(0.8)^3(0.2)^2$;
$5000\ {}^5C_4(0.8)^4(0.2)$; $5000(0.8)^5$

7.22 Given p = probability that he is capable of making a distinction $= 0.75$; $q = 1 - p = 0.25$

(i) $P(x < 5) = 1 - P(x \geq 5) = 1 - [{}^6C_5(0.75)^5(0.25) + {}^6C_6(0.75)^6] = 1 - 0.534 = 0.466$

(ii) $P(x \geq 5) = 0.534$

7.5.2 Poisson Probability Distribution

Poisson (named after the French mathematician S. Poisson, 1781–1840) probability distribution is widely used for a discrete random variable such as (i) number of telephone calls per hour coming into the switchboard, (ii) number of traffic accidents per week in a city/state, (iii) number of patients arriving at a hospital every hour, (iv) number of cars waiting for service in a workshop, and so on.

The Poisson process measures the number of occurrences of an outcome of a discrete random variable in a *predetermined time interval, space or volume*, for which an *average number* of occurrences is known or can be determined. The Poisson probability distribution is an approximation to a binomial distribution when the probability of success, p is very small and n is large, so that $\mu = np$ is small, preferably $np > 7$. It is often called the '*law of improbable*' and implies that the probability, p, of a particular outcomes occurrence is very small.

Poisson Distribution: A discrete probability distribution in which the probability of occurrence of an outcome within a very small time period is very small.

As mentioned above **Poisson distribution** occurs in business situations in which there are only few successes in an interval of time against a large number of failures or vice-versa and has single independent outcomes that are mutually exclusive. Because of this, the probability of success, p is very small in relation to the number of trials n, so only the probability of success is considered.

Conditions for Poisson Process

The use of Poisson distribution to compute the probability of the occurrence of an outcome during a specific time period is based on the following conditions:

- The outcomes in any interval of time occur randomly and independently of one another.
- The probability of occurrence of an outcome in a small interval of time is proportional to the length of the interval but is independent of the specific time interval.
- The probability of occurrence of more than one outcome in a small interval of time is negligible.
- The average number of occurrence of outcomes is constant in equal intervals of time.

Poisson Probability Function

If an experiment is repeated a large number of times, say n, then probability, p, of occurrence of an outcome of interest (i.e., success) in each trial is very small. The average number of times that an outcome occurs in a certain period of time, $\lambda = np$ is also small. Under these conditions, the binomial probability function

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

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{r-1}{n}\right)\frac{\lambda^r}{r!}\left(1-\frac{\lambda}{n}\right)^{n-r}; np = \lambda \text{ or } p = \lambda/n$$

tends to $\frac{\lambda^r}{r!}e^{-\lambda}$ for a fixed value of r. Thus, the Poisson probability distribution which approximates the binomial distribution is defined by the following probability function:

$$P(x = r) = \frac{\lambda^r e^{-\lambda}}{r!}, r = 0, 1, 2, \ldots \qquad (7\text{-}3)$$

where $e = 2.7183$.

Characteristics of Poisson Distribution

Since Poisson probability distribution is specified by a process rate λ and the time period t, its mean and variance are identical and expressed in terms of the parameters n and p as shown below:

- ***Arithmetic mean***, $\mu = E(x)$ of Poisson distribution is given by

$$\mu = \sum x P(x) = \sum x \frac{e^{-\lambda}\lambda^x}{x!}, x = 1, 2, 3, \ldots \text{ and } x P(x) = 0 \text{ for } x = 0$$

$$= \lambda e^{-\lambda} + \lambda^2 e^{-\lambda} + \frac{\lambda^3 e^{-\lambda}}{2!} + \ldots + \frac{\lambda^r e^{-\lambda}}{(x-1)!} + \ldots$$

$$= \lambda e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^{r-1}}{(x-1)!} + \ldots\right] = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Thus, the mean of the distribution is $\mu = \lambda = np$.

- ***Variance*** σ^2 of Poisson distribution is given by

$$\sigma^2 = E(x^2) - [E(x)]^2 = E(x^2) - \lambda^2$$

$$= \sum x^2 \frac{e^{-\lambda}\lambda^x}{x!} - \lambda^2 = e^{-\lambda}\sum \frac{x(x-1)+x}{x!}\lambda^x - \lambda^2$$

$$= \lambda^2 e^{-\lambda}\sum \frac{\lambda^{x-2}}{(x-2)!} + \lambda e^{-\lambda}\sum \frac{\lambda^{x-1}}{(x-1)!} - \lambda^2$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Thus, the variance of the distribution is $\sigma^2 = \lambda = np$.

The central moments of Poisson distribution can also be determined by the following recursion relation:

$$\mu_r = E(x-\lambda)^r = \sum (x-\lambda)^r e^{-\lambda}\frac{\lambda^x}{x!}$$

Differentially μ_r with respect to λ, we have

$$\frac{d\mu_r}{d\lambda} = -r\mu_{r-1} + \frac{\mu_{r+1}}{\lambda} \text{ or } \mu_{r+1} = \lambda\left[r\mu_{r-1} + \frac{d\mu_r}{d\lambda}\right] \quad (7\text{-}4)$$

Substituting $\mu_0 = 1$ and $\mu_1 = 0$ and putting $r = 1, 2$ and 3 in equation (7-4), we have

$$\mu_2 = \mu_3 = \lambda$$
$$\mu_4 = \lambda + 3\lambda^2$$

so that $\gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}$ and $\gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}$

The chances for *more than one event to occur* during a short interval of time are less. The probability that exactly one event will occur in such an interval is approximately λ times its duration.

If λ is not an integer and $m = [\lambda]$, the largest integer contained in it, then distribution has unique mode. But if λ is an integer, the distribution would be bimodal.

The typical application of Poisson distribution is for analysing queuing (or waiting line) problems in which customers during an interval of time arrive independently and the average or mean number of arrivals (or events) is proportional to the length of the time period.

Fitting a Poisson Distribution

Poisson distribution can be fitted to the observed values in the data set by obtaining values of λ and calculating the probability of zero occurrence. Other probabilities can be calculated by the recurrence relation as follows:

$$P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

$$P(r+1) = \frac{e^{-\lambda}\lambda^{r+1}}{(r+1)!}$$

or

$$\frac{f(x+1)}{f(r)} = \frac{\lambda}{r+1} \text{ or } P(r+1) = \frac{\lambda}{(r+1)} P(r);\ r = 0, 1, 2, \ldots$$

Thus, for $r = 0$, $P(1) = \lambda P(0)$,

For $r = 1$, $P(2) = \frac{\lambda}{2} P(1) = \frac{\lambda^2}{2} P(0)$

and so on, where $P(0) = e^{-\lambda}$.

After obtaining the probability for each of the random variable values, multiply them by N (total frequency) to get the expected frequency for the respective values.

Example 7.15: What probability model is appropriate to describe a situation where 100 misprints are distributed randomly throughout the 100 pages of a book? For this model, what is the probability that a page observed at random will contain at least three misprints?

Solution: Since 100 misprints are distributed randomly throughout the 100 pages of a book, therefore on an average there is only one mistake on a page. This means the probability of there being a misprint, $p = 1/100$, is very small and the number of words, n, in 100 pages are vary large. Hence, Poisson distribution is best suited in this case.

Average number of misprints in one page, $\lambda = np = 100 \times (1/100) = 1$. Therefore, $e^{-\lambda} = e^{-1} = 0.3679$.

Probability of at least three misprints in a page is

$$P(x \geq 3) = 1 - P(x < 3) = 1 - \{P(x = 0) + P(x = 1) + P(x = 2)\}$$

$$= 1 - [e^{-\lambda} + \lambda e^{-\lambda} + \frac{1}{2!}\lambda^2 e^{-\lambda}]$$

$$= 1 - \left\{e^{-1} + e^{-1} + \frac{e^{-1}}{2!}\right\} = 1 - 2.5\, e^{-1}$$

$$= 1 - 2.5\,(0.3679) = 0.0802$$

Example 7.16: A new automated production process has had an average of 1.5 breakdowns per day. Because of the cost associated with a breakdown, management is concerned about the possibility of having three or more breakdowns during a day. Assume that breakdowns occur randomly that the probability of a breakdown is the same for any two time intervals of equal length, and that breakdowns in one period are independent of breakdowns in other periods. What is the probability of having three or more breakdowns during a day? *[HP Univ., MBA, 2005; Kumaon Univ., 2008]*

Solution: Given that $\lambda = np = 1.5$ breakdowns per day. Thus, probability of having three or more breakdowns during a day is given by

$$P(x \geq 3) = 1 - P(x < 3) = 1 - [P(x = 0) + P(x = 1) + P(x = 2)]$$

$$= 1 - \left[\frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!}\right]$$

$$= 1 - e^{-\lambda}\left[1 + \lambda + \frac{1}{2}\lambda^2\right] = 1 - 0.2231\left[1 + 1.5 + \frac{1}{2}(1.5)^2\right]$$

$$= 1 - 0.2231\,(3.625) = 1 - 0.8088 = 0.1912$$

Example 7.17: Suppose a life insurance company insures the lives of 5000 persons aged 42. If studies show the probability that any 42-years old person will die in a given year to be 0.001, find the probability that the company will have to pay at least two claims during a given year.

Solution: Given that $n = 5000$, $p = 0.001$, so $\lambda = np = 5000 \times 0.001 = 5$. Thus, the probability that the company will have to pay at least 2 claims during a given year is given by

$$P(x \geq 2) = 1 - P(x < 2) = 1 - [P(x = 0) + P(x = 1)]$$

$$= 1 - [e^{-\lambda} + \lambda e^{-\lambda}] = 1 - [e^{-5} + 5e^{-5}] = 1 - 6e^{-5}$$

$$= 1 - 6 \times 0.0067 = 0.9598$$

Example 7.18: A manufacturer who produces medicine bottles finds that 0.1 per cent of the bottles are defective. The bottles are packed in boxes containing 500 bottles. A drug manufacturer buys 100 boxes from the producer of bottles. Using Poisson distribution, find how many boxes will contain:

(i) no defectives.

(ii) at least two defectives. *[Delhi Univ., MBA, 2001, 2006]*

Solution: Given that $p = 1$ per cent $= 0.001$, $n = 500$, $\lambda = np = 500 \times 0.001 = 0.5$

(i) $P[x = 0] = e^{-\lambda} = e^{-0.5} = 0.6065$

Therefore, the required number of boxes are $0.6065 \times 100 = 61$ (approx.)

(ii) $P(x > 2) = 1 - P(x \le 1) = 1 - [P(x = 0) + P(x = 1)]$

$= 1 - [e^{\lambda} + \lambda e^{-\lambda}] = 1 - [0.6065 + 0.5(0.6065)]$

$= 1 - 0.6065(1.5) = 1 - 0.90975 = 0.09025.$

Therefore, the required number of boxes are $100 \times 0.09025 = 10$ (approx.)

Example 7.19: In a town, 10 accidents took place in a span of 50 days. Assuming that the number of accidents per day follows the Possion distribution, find the probability that there will be three or more accidents in a day. *[Coimbatore Univ., MBA, 2005]*

Solution: The average number of accidents per day are $10/50 = 0.2$. Thus,

$$\begin{aligned} P\text{ (3 or more accidents)} &= 1 - P\text{ (2 or less accidents)} \\ &= 1 - [P(0) + P(1) + P(2)] \\ &= 1 - \left[e^{-2} + e^{-2} \times 0.2 + \frac{e^{-2} \times (0.2)^2}{2!} \right] \\ &= 1 - e^{-2}\,[1 + 0.2 + 0.02] = 1 - e^{-2} \times 1.22 \\ &= 1 - 0.8187 \times 1.22 = 1 - 0.999 = 0.001. \end{aligned}$$

Example 7.20: When the first proof of 200 pages of a book of 5000 pages was read, the distribution of printing mistakes found are shown in the table. Fit a Poisson distribution to the frequency distribution of printing mistakes. Estimate the total cost of correcting the whole book by using the information provided in the table below:

Number of Misprints per Page	*Frequency*	*Cost of Detection and Correction per Page (₹)*
0	113	1.00
1	62	1.50
2	20	2.50
3	3	3.00
4	1	3.50
5	1	4.00

[MD Univ., MBA, 2005]

Solution: Calculations required to fit a Poisson distribution are shown in Table 7.3:

Table 7.5 Calculations for Poisson Distribution

Number of Mistakes per Page (x)	*Frequency (f)*	*f x*
0	113	0
1	62	62
2	20	40
3	3	9
4	1	4
5	1	5
	200	120

$$\bar{x}\,(=\lambda) = \frac{1}{n}\sum fx = \frac{120}{200} = 0.6$$

Expected number of misprints are as follows:

$$\begin{aligned} nP(x = 0) &= ne^{-m} = 200e^{-0.6} = 200 \times 0.5488 = 109.76 \\ nP(x = 1) &= nP(x = 0) \times \lambda = 109.76 \times 0.6 = 65.856 \\ nP(x = 2) &= nP(x = 0) \times \frac{\lambda}{2} = 65.856 \times \frac{0.6}{2} = 19.756 \\ nP(x = 3) &= nP(x = 1) \times \frac{\lambda}{3} = 19.757 \times \frac{0.6}{3} = 3.951 \end{aligned}$$

$$nP(x = 4) = nP(x = 3) \times \frac{\lambda}{4} = 3.951 \times \frac{0.6}{4} = 0.593$$

$$nP(x = 5) = nP(x = 4) \times \frac{\lambda}{5} = 0.5927 \times \frac{0.6}{5} = 0.071.$$

The expected number of pages containing mistakes are as follows:

Mistakes :	0	1	2	3	4	5
Pages expected :	109.76	65.856	19.756	3.951	0.593	0.071

Calculations for total cost of correcting the first proof of the book are shown in the Table 7.6.

Table 7.6

Number of Misprints per Page	*Rate per Page (x)*	*Number of Pages (f)*	*fx*
0	1.00	109.76	109.760
1	1.50	65.856	98.784
2	2.50	19.756	49.390
3	3.00	3.951	11.853
4	3.50	0.593	2.075
5	4.00	0.071	0.284
			272.146

Hence, the total cost shall be ₹272.146.

Example 7.21: The following table shows the number of customers returning the products in a marketing territory. The data is for 100 stores;

Number of returns :	0	1	2	3	4	5	6
Customers :	4	14	23	23	18	9	9

Fit a Poisson distribution. [*Lucknow Univ., MBA, 2006*]

Solution: Calculations required to fit a Poisson distribution are shown in Table 7.7:

Table 7.7 Calculation for Poisson Distribution

Number of Returns (x)	*Customers (f)*	*fx*
0	4	0
1	14	14
2	23	46
3	23	69
4	18	72
5	9	45
6	9	54
	n = 100	300

$$\bar{x}\,(=\lambda) = \frac{1}{n}\sum fx = \frac{300}{100} = 3.$$

Hence, expected frequencies (customers returning products) are as follows:

$$P(x=0) = e^{-3} = 0.0498$$

$$nP\,(x=0) = 100 \times 0.0498 = 4.98$$

$$nP(x=1) = nP(x = 0)\times\lambda = 4.98\times 3 = 14.94$$

$$nP(x=2) = nP(x=1) \times \frac{\lambda}{2} = 14.94 \times \frac{3}{2} = 22.41$$

$$nP(x=3) = nP(x=2) \times \frac{\lambda}{3} = 22.41 \times \frac{3}{3} = 22.41$$

$$nP(x=4) = nP(x=3)\times\frac{\lambda}{4} = 22.41\times\frac{3}{4} = 16.80$$

$$nP(x=5) = nP(x=4)\times\frac{\lambda}{5} = 16.80\times\frac{3}{5} = 10.08$$

$$nP(x=6) = nP(x=5)\times\frac{\lambda}{6} = 10.08\times\frac{3}{6} = 5.04.$$

The expected number of customers returning the products in a marketing territory are:

Number of returns :	0	1	2	3	4	5	6
Expected customers :	4.98	11.94	22.41	22.41	16.80	10.08	5.04

Example 7.22: The following table gives the number of days in a 50-day period during which automobile accidents occurred in a city:

No. of accidents :	0	1	2	3	4
No. of days :	21	18	7	3	1

Fit a Poisson distribution to the data. *[Sukhadia Univ., MBA, 2002; Kumaon Univ., MBA,2000]*

Solution: Calculations for fitting of Poisson distribution are shown in the Table 7.8.

Table 7.8 Calculations for Poisson Distribution

Number of Accidents (x)	*Number of Days (f)*	*fx*
0	21	0
1	18	18
2	7	14
3	3	09
4	1	04
	n = 50	45

Thus, $$\bar{x}\text{(or }\lambda) = \frac{1}{n}\sum fx = \frac{45}{50} = 0.9$$

and $$P(x=0) = e^{-\lambda} = e^{-0.9} = 0.4066$$

$$P(x=1) = \lambda P(x=0) = 0.9(0.4066) = 0.3659$$

$$P(x=2) = \frac{\lambda}{2}P(x=1) = \frac{0.9}{2}(0.3659) = 0.1647$$

$$P(x=3) = \frac{\lambda}{3}P(x=2) = \frac{0.9}{3}(0.1647) = 0.0494$$

$$P(x=4) = \frac{\lambda}{4}P(x=3) = \frac{0.9}{4}(0.0494) = 0.0111$$

In order to fit a Poisson distribution, we shall multiply each of these values by n = 50 (total frequencies). Hence the expected frequencies are

Accidents :	0	1	2	3	4
Days :	0.4066 × 50 = 20.33	0.3659 × 50 = 18.30	0.1647 × 50 = 8.23	0.0494 × 50 = 2.47	0.0111 × 50 = 0.56

7.5.3 Negative Binomial Probability Distribution

All conditions of binomial distribution are also applicable to the negative binomial distribution except that it describes the number of trials likely to be required to obtain a fixed number of successes. For example, suppose a percentage p of individuals in the population is sampled until exactly r individuals with the certain characteristic are found. The number of individuals in excess of r that are observed or sampled has a negative binomial distribution.

The probability distribution function of the negative binomial distribution is obtained by considering an infinite series of Bernoulli trials with probability of success p of an event on an individual trial. If trials are repeated r times until an event of interest occurs, then the probability that at least m trials will be required to get the event r times (successes) is given by

$$P(m, r, p) = \text{Probability that an event occurs } (r - 1) \text{ times in the first } m - 1 \text{ trials} \times \text{Probability that the event of interest occurs in the } m\text{th trial}$$

$$= {}^{m-1}C_{r-1}\, p^{r-1}\, q^{m-r} \times p = {}^{m-1}C_{r-1}\, p^{r}\, q_{m-r}, \quad m = r, r+1, \ldots \tag{7-5}$$

where $r \geq 1$ is a fixed integer.

A random variable having a negative binomial distribution is also referred to as a discrete waiting time random variable. In terms of number of failures, it represents how long one waits for the rth success.

The mean and variance of this distribution are given by

$$\text{Mean, } \mu = \frac{r}{p}, \text{ Variance, } \sigma^2 = \frac{rq}{p^2}$$

Example 7.23: A market research agency that conducts interviews by telephone has found from past experience that there is a 0.40 probability that a call made between 2.30 P.M. and 5.30 P.M. will be answered. Assuming a Bernoullian process:

(a) Calculate the probability that an interviewer's 10th answer comes on his 20th call and that he will receive the first answer on his 3rd call.

(b) What is the expected number of calls necessary to obtain seven answers?

Solution: Let answer to a call be considered 'success'. Then $p = 0.40$

(a) P(10th answer comes on 20th call)

$$= {}^{m-1}C_{r-1}\, p^{r}\, q^{m-r}, m = r, r+1, \ldots$$

$$= {}^{19}C_9\, (0.4)^{10}\, (0.6)^{10};\ m = 20 \text{ and } r = 10$$

$$= 0.058$$

$$\text{P(First answer on 3rd call)} = {}^{2}C_0 (0.4)^1\, (0.6)^2 = 0.144$$

(b) Expected number of calls for 7 answers, $\mu = r/p = 7/0.4 = 17.5 \cong 18$ calls

7.5.4 Multinomial Probability Distribution

The binomial distribution discussed earlier is associated with a sequence of n independent repeated Bernoulli trials, each resulting in only two outcomes, and one of the possible outcomes is called success, while multinomial distribution is associated with independent repeated trials that generalize from Bernoulli trials each resulting in two to k outcomes.

Suppose a single trial of an experiment results in only one of the k possible outcomes $O_1, O_2, \ldots, O_k$ with respective probabilities $p_1, p_2, \ldots, p_k$, and the experiment is repeated n times independently. The probability that out of these n trials outcome O_1 occurs x_1 times, O_2 occurs x_2 times and so on is given by the following discrete density function:

$$P(x_1, x_2, \ldots, x_k) = \frac{n!}{x_1!\, x_2! \ldots x_k!}\left[p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}\right] \tag{7-6}$$

where $x_1 + x_2 + \ldots + x_k = n$.

Example 7.24: In a factory producing certain items, 30 per cent of the items produced have no defect, 40 per cent have one defect, and 30 per cent have two defects. A random sample of 8 items is taken from a day's output. Find the probability that it will contain 2 items with no defect, 3 items with one defect and 3 items with two defects.

Solution: We know from the data that $n = 8, p_1 = 0.30, p_2 = 0.40$ and $p_3 = 0.30; x_1 = 2, x_2 = 3, x_3 = 3$. Thus, the required probability is given by

$$P(x_1 = 2, x_2 = 3, x_3 = 3) = \frac{8!}{2!\,3!\,3!}\,[(0.30)^2\,(0.40)^2\,(0.30)^3]$$

$$= 0.0871$$

7.5.5 Hyper-geometric Probability Distribution

For a binomial distribution to be applied, the probability of a success or failure must remain the same for each trial. This is possible only when the number of elements in the population are large relative to the number in the sample, where probability of getting a success on a single trial is equal to the proportion p of successes in the population. However, if the number of element in the population are small in relation to the sample size, i.e. $n/N \geq 0.5$, the probability of a success in a given trial is dependent upon the outcomes of preceding trials. Then the number r of successes follows hyper-geometric probability distribution. Thus, hyper-geometric probability distribution is similar to binomial distribution where probability of success may be different from trial to trial. When sampling is done *without replacement* from a finite population, the Bernoulli process does not apply because there is a systematic change in the probability of success in the reduced size of population.

Let N be the size of population and out of N, m be the total number of elements having a certain characteristic (called success) and the remaining $N - m$ do not have it, such that $p + q = 1$. Suppose a sample of size n is drawn at random without replacement. Then in a random sample of size n, the probability of exactly r successes when values of r depend on N, p and n are given by

$$P(x = r) = \frac{{}^{m}C_{r}\,{}^{N-m}C_{n-r}}{{}^{N}C_{n}};\ r = 0, 1, 2, \ldots, n;\ \text{and}\ 0 \leq r \leq m$$

This probability mass function is called *hyper-geometric probability distribution*.

The mean and variance of a hyper-geometric distribution are

$$\text{Mean} = n\left(\frac{m}{N}\right) \text{ and Variance} = n\left(\frac{m}{N}\right)\left(\frac{N-m}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Example 7.25: Suppose the HRD manager randomly selects 3 individuals from a group of 10 employees for a special assignment. Assuming that 4 of the employees were assigned to a similar assignment previously, determine the probability that exactly two of the three employees have had previous experience.

Solution: We know from the data that $N = 10, n = 3, r = 2, m = 4$, and $N - m = 6$. Thus, the required probability is given by

$$P(x = r \mid N, m, N-m) = \frac{{}^{m}C_{r}\,{}^{N-m}C_{n-r}}{{}^{N}C_{n}} = \frac{{}^{4}C_{2}\,{}^{10-4}C_{3-2}}{{}^{10}C_{3}}$$

$$= \frac{{}^{4}C_{2}\,{}^{6}C_{1}}{{}^{10}C_{3}} = \frac{\left(\frac{4!}{2!\,2!}\right)\left(\frac{6!}{1!5!}\right)}{\left(\frac{10!}{3!7!}\right)} = \frac{36}{120} = 0.30$$

Example 7.26: Suppose a particular industrial product is shipped in lots of 20. To determine whether an item is defective a sample of 5 items from each lot is drawn. A lot is rejected if more than one defective item is observed. (If the lot is rejected, each item in the lot is then tested). If a lot contains four defectives, what is the probability that it will be accepted?

Solution: Let r be the number of defectives in the sample size $n = 5$. Given that, $N = 20$, $m = 4$, and $N - m = 16$. Then,

$$P(\text{accept the lot}) = P(x \le 1) = P(x = 0) + P(x = 1)$$

$$= \frac{{}^{4}C_0 \times {}^{16}C_5}{{}^{20}C_5} + \frac{{}^{4}C_1 \times {}^{16}C_4}{{}^{20}C_5} = \frac{\frac{4!}{0!4!} \times \frac{16!}{5!11!}}{\frac{20!}{5!15!}} + \frac{\frac{4!}{1!3!} \times \frac{16!}{4!12!}}{\frac{20!}{5!15!}}$$

$$= \frac{91}{323} + \frac{455}{969} = 0.2817 + 0.4696 = 0.7513$$

Conceptual Questions 7C

13. If x has a Poisson distribution with parameter λ, then show that $E(x)$ and $V(x) = \lambda$. Further, show that the Poisson distribution is a limiting form of the binomial distribution.

14. What is Poisson distribution? Point out its role in business decision-making. Under what conditions will it tend to become a binomial distribution?

[Kumaon Univ., MBA, 1998]

15. When can Poisson distribution be a reasonable approximation of the binomial?

[Delhi Univ., M.Com., 1999]

16. Discuss the distinctive features of Poisson distribution. When does a binomial distribution tend to become a Poisson distribution?

17. Under what conditions is the Poisson probability distribution appropriate? How are its mean and variance calculated?

18. What is negative binomial distribution? Distinguish the relationship between the binomial and negative binomial distributions.

19. What is hyper-geometric distribution? Explain its properties.

Self-practice Problems 7C

7.23 In a certain factory manufacturing razor blades, there is a small chance of 1/150 for any blade to be defective. The blades are placed in packets, each containing 10 blades. Using the Poisson distribution, calculate the approximate number of packets containing not more than 2 defective blades in a consignment of 10,000 packets.

7.24 The distribution of typing mistakes committed by a typist is given below. Assuming a Poisson distribution, find out the expected frequencies:

No. of mistakes per page :	0	1	2	3	4	5
No. of pages :	142	156	69	27	5	1

[Rohilkhand Univ., MBA, 2006]

7.25 Find the probability that at most 5 defective bolts will be found in a box of 200 bolts if it is known that 2 per cent of such bolts are expected to be defective [you may take the distribution to be Poisson; $e^{-4} = 0.0183$].

7.26 On an average, one in 400 items is defective. If the items are packed in boxes of 100, what is the probability that any given box of items will contain: (i) no defectives; (ii) less than two defectives; (iii) one or more defectives; and (iv) more than three defectives

[Delhi Univ., MBA, 2000]

7.27 It is given that 30 per cent of electric bulbs manufactured by a company are defective. Find the probability that a sample of 100 bulbs will contain (i) no defective, and (ii) exactly one defective.

7.28 One-fifth per cent of the blades produced by a blade manufacturing factory turn out 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 1,00,000 packets. *[Delhi Univ., MBA, 2003, 2008]*

7.29 A factory produces blades in packets of 10. The probability of a blade to be defective is 0.2 per cent. Find the number of packets having two defective blades in a consignment of 10,000 packets.

7.30 In a certain factory manufacturing razor blades, there is small chance 1/50 for any blade to be defective. The blades are placed in packets, each containing 10 blades. Using an appropriate probability distribution, calculate the approximate number of packets containing not more than 2 defective blades in a consignment of 10,000 packets.

7.31 Suppose that a manufactured product has 2 defects per unit of product inspected. Using Poisson distribution, calculate the probabilities of finding a product without any defect, 3 defects, and 4 defects. (Given $e^{-2} = 0.135$) *[Madurai Univ., M.Com., 2004]*

7.32 A distributor received a shipment of 12 TV sets. Shortly after this shipment was received, the manufacturer informed that he had inadvertently shipped 3 defective sets. The distributor decided to test 4 sets randomly selected out of 12 sets received.

(a) What is the probability that neither of the 4 sets tested was defective?

(b) What is the mean and variance of defective sets.

7.33 Suppose a population contains 10 elements, 6 of which are defective. A sample of 3 elements is selected. What is the probability that exactly 2 are defective?

7.34 A transport company has a fleet of 15 trucks, used mainly to deliver fruits to wholesale market. Suppose 6 of the 15 trucks have brake problems. Five trucks were selected at random to be tested. What is the probability that 2 of those tested trucks have defective brakes?

7.35 A company has five applicants for two positions: two women and three men. Suppose that the five applicants are equally qualified and that no preference is given for choosing either gender. If r equal the number of women chosen to fill the two positions, then what is the probability distribution of r. Also, determine the mean and variance of this distribution.

Hints and Answers

7.23 Given that N = 10,000, $p = 1/50$, $n = 10$, $\lambda = np = 10\times(1/50) = 0.2$

$P(x = 0) = e^{-\lambda} = e^{-0.2} = 0.8187$ (from the table)

$NP(x = 0) = 0.8187 \times 10{,}000 = 8187$

$NP(x = 1) = NP(x = 0) \times \lambda = 8187 \times 0.2 = 1637.4$

$NP(x = 2) = NP(x = 1) \times \lambda/2 = 1637.4 \times (0.2/2)$
$= 163.74$

The approximate number of packets containing not more than 2 defective blades in a consignment of 10,000 packets is: 10,000 – (8187+163.74+163.74) = (10,000 – 9988.14) = 11.86 or 12.

7.24 $\lambda = \frac{1}{n}\sum fx = 400/400 = 1,$

$P(x = 0) = e^{-\lambda} = 0.3679$

$NP(x = 0) = 147.16$

$NP(x = 1) = NP(x = 0)\lambda = 147.16$

$NP(x = 2) = NP(x = 1)\frac{\lambda}{2} = 73.58$

$NP(x = 3) = NP(x = 2)\frac{\lambda}{3} = 24.53$

$NP(x = 4) = NP(x = 3)\frac{\lambda}{4} = 6.13$

$NP(x = 5) = NP(x = 4)\frac{\lambda}{5} = 1.23$

Expected frequencies as per the distribution are:

No. of mistakes per page	:	0	1	2	3	4	5
No. of pages	:	147	147	74	25	6	1

7.25 P(defective bolt) = 2 per cent = 0.02. Given $n = 200$, so $\lambda = np = 200 \times 0.02 = 4$

$$P(0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-4} = 0.0183$$

$$P(x \le 5) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5)$$

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

$= 0.0183 \times (643/15) = 0.7844.$

7.27 $\lambda = np = 100 \times 0.30 = 3$

$P(x = 0) = e^{-\lambda} = e^{-3} = 0.05;$

$P(x = 1) = \lambda P(x = 0) = 3 \times 0.03 = 0.15$

7.28 Given $n = 10$, $p = 1/500$, $\lambda = np = 10/500 = 0.02$

(i) $P(x = 0) = e^{-\lambda} = e^{-0.02} = 0.9802$

$NP(x = 0) = 1{,}00{,}000 \times 0.9802$
$= 98020$ packets

(ii) $P(x = 1) = \lambda P(x = 0) = \lambda e^{-\lambda}$
$= 0.02 \times 0.9802$
$= 0.019604$

$NP(x = 1) = 1{,}00{,}000 \times 0.019604$
$= 1960$ packets

(iii) $P(x = 2) = \frac{\lambda^2}{2}P(x = 0) = \frac{(0.02)^2}{2}\times 0.9802$
$= 0.00019604$

$NP(x = 2) = 1{,}00{,}000 \times 0.00019604$
$= 19.60 \cong 20$ packets

7.29 Given $n = 10$, $p = 0.002$, $\lambda = np = 10 \times 0.002 = 0.02.$

$$P(x = 2) = \frac{e^{-\lambda}\lambda^2}{2!} = \frac{e^{-0.02}(0.02)^2}{2!}$$
$= 0.000196$

The required number of packets having two defective blades each in a consignment of 10,000 packets $= 10{,}000\times 0.000196 \cong 2.$

7.30 Given N = 10,000, $p = 1/50$ and $n = 10$.

Thus $\lambda = np = 0.20$ and

$NP(x = 0) = 10{,}000\, e^{-\lambda} = 10{,}000\, e^{-0.20} = 8187;$

$P(x = 1) = NP(x = 0) \times \lambda = 8187 \times 0.2 = 1637.4.$

$NP(x = 2) = NP(x = 1) \times \frac{\lambda}{2} = 1637.4 \times \frac{0.2}{2}$
$= 163.74$

The approximate number of packets containing not more than 2 defective blades in a consignment of 10,000 packets will be: 10,000 – (8187 + 1637.40 + 163.74) = 12 approx.

7.31 Given average number of defects, $\lambda = 2$.

$$P(x = 0) = e^{-\lambda} = e^{-2} = 0.135;$$

$$P(x = 1) = P(x = 0) \times \lambda = 0.135 \times 2 = 0.27$$

$$P(x = 2) = P(x = 1) \times \frac{\lambda}{2} = 0.27 \times \frac{2}{2} = 0.27$$

$$P(x = 3) = P(x = 2) \times \frac{\lambda}{3} = 0.27 \times \frac{2}{3} = 0.18$$

$$P(x = 4) = P(x = 3) \times \frac{\lambda}{4} = 0.18 \times \frac{2}{4} = 0.09$$

7.32 Given $N = 12, n = 4, m = 3$ and $N - m = 9$.

(a) $$P(x = 0) = \frac{{}^{3}C_0 \times {}^{9}C_4}{{}^{12}C_4} = \frac{\frac{3!}{0!3!} \times \frac{9!}{4!5!}}{\frac{12!}{4!8!}}$$

$$= \frac{1 \times 126}{495} = \frac{14}{55}$$

(b) $$\text{Mean, } \mu = n\left(\frac{m}{N}\right) = 4\left(\frac{3}{12}\right) = 1$$

$$\text{Variance, } \sigma^2 = n\left(\frac{m}{N}\right)\left(\frac{N-m}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$= 4\left(\frac{3}{12}\right)\left(\frac{9}{12}\right)\left(\frac{8}{11}\right) = 0.5455$$

7.33 $$P(x = 2) = \frac{{}^{6}C_2 \times {}^{4}C_1}{{}^{10}C_3} = \frac{15 \times 4}{120} = 0.50$$

7.34 $$P(x = 2) = \frac{{}^{9}C_3 \times {}^{6}C_2}{{}^{15}C_5} = \frac{84 \times 15}{3003} = 0.4196$$

7.35 Given $N = 5, n = 2, m = 2, N - m = 3$

$$P(x = r) = \frac{{}^{m}C_r \times {}^{N-m}C_{n-r}}{{}^{N}C_n} = \frac{{}^{2}C_r \times {}^{3}C_{2-r}}{{}^{5}C_2};\ r = 0, 1, 2$$

$$\text{Mean, } \mu = 2\left(\frac{2}{5}\right) = 0.8;$$

$$\text{Variance} = 2\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)\left(\frac{3}{4}\right) = 0.6$$

7.6 CONTINUOUS PROBABILITY DISTRIBUTIONS

Since continuous random variables such as height, time, weight, monetary values, length of life of a particular product, etc. can take large number of both integer and non-integer values. The sum of the probability to each of these values is no longer sum to 1.

Unlike discrete random variables, continuous random variables do not have probability distribution functions specifying the exact probabilities of their specified values. Instead, probability distribution is created by distributing one unit of probability along the real line. Such a distribution (also called *probability density function)* determines probabilities that the random variable falls into a specified interval of values.

Certain characteristics of probability density function for the continuous random variable, x, are follows:

(i) Area under a continuous probability distribution is equal to 1.

(ii) Probability $P(a \le x \le b)$ of random variable, x, value will fall in an interval from a to b is equal to the area under the *probability density function curve* between the points (values) a and b.

Since nature follows a predictable pattern for many kinds of measurements, therefore most numerical values of a random variable are spread around the center. A frequency distribution of values of random variable observed in nature which follows this pattern is approximately bell shaped. Thus, such distribution of measurements is called a **normal curve (or distribution).**

Normal Distribution: A continuous probability distribution in which the mean of the distribution lies at the center of the curve and the curve is symmetrical around a vertical line erected at the mean. The tails of the curve extend indefinitely parallel to the horizontal axis.

German mathematician Karl Friedrich Gauss developed the concept of normal distribution (also known as *Gaussian distribution*). Normal distribution is used to study a continuous phenomenon or process such as daily changes in the stock market index, frequency of arrivals of customers at a bank, frequency of telephone calls into a switch board, customer servicing times and so on.

7.6.1 Normal Probability Distribution Function

The formula for normal probability distribution is as follows:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{(-1/2)[(x-\mu)/\sigma]^2},\ -\infty < x < \infty \qquad (7\text{-}7)$$

where π is constant 3.1416, e is constant 2.7183, μ is the mean of the normal distribution, σ is standard of normal distribution and $f(x)$=relative frequencies (height of the curve) within which values of random variable x fall.

The graph of a normal probability distribution with mean μ and standard deviation σ is shown in Fig. 7.7. The distribution is symmetric about its mean μ that falls at the centre of curve. Since the total area under the normal probability distribution is equal to 1, the symmetry implies that the area on either side of μ is 50 per cent or 0.5. The *shape* of the distribution is determined by μ and σ values.

In symbols, if a random variable x follows normal probability distribution with mean μ and standard deviation σ, then it is also expressed as $x \sim N(\mu, \sigma)$.

Characteristics of Normal Probability Distribution

The shape of normal distribution varies according to the value of mean, μ and/or standard deviation σ. Larger the value of the standard deviation σ, the wider and flatter is the normal curve showing more variability in the data. Thus, standard deviation σ determines the range of values that any random variable is likely to assume. Figure 7.6(a) shows three normal distributions with different values of the mean μ and a fixed standard deviation σ while in Fig. 7.6(b) normal distributions are shown with different values of the standard deviation σ and a fixed mean μ.

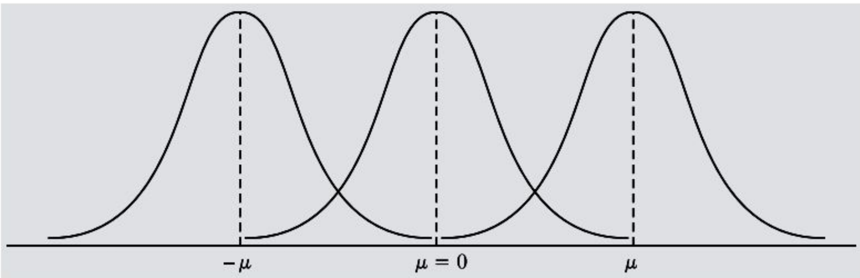


Figure 7.6 (a)
Normal Distributions with Different Mean Values but Fixed Standard Deviation

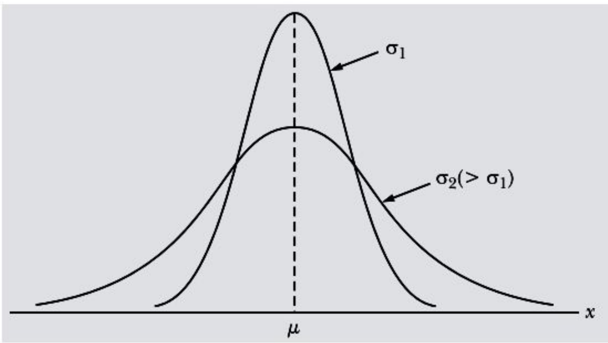


Figure 7.6 (b)
Normal Distributions with Fixed Mean and Variable Standard Deviation

From Figs 7.6(a) and 7.6(b), the following characteristics of a normal distribution and its density function may be derived:

(i) For every pair of values of μ and σ, the normal probability density function curve is bell shaped and symmetric. The mean μ determines the *central location* of the normal distribution while standard deviation σ determines its *spread*.

(ii) The normal curve is symmetrical around a vertical line erected at the mean μ with respect to the area under it, i.e., 50 per cent of the area of the curve lies on both sides of the mean, μ. This implies that the probability of any random variable whose value is above or below the mean will be same. Thus, for any normal random variable x, $P(x \leq \mu) = P(x \geq \mu) = 0.50$

(iii) The values of mean, median and mode for the normal distribution are equal because the highest value of the probability density function occurs when value of a random variable, $x = \mu$.

(iv) The two tails of the normal curve extend to infinity in both directions and never touch the horizontal axis.

(v) The mean of the normal distribution may be negative, zero or positive as shown in Fig. 7.6(a).

(vi) The area under the normal curve represents probabilities for the normal random variable, and therefore, the total area under the curve for the normal probability distribution is 1.

Standard Normal Probability Distribution: A normal probability distribution with mean equal to zero and standard deviation equal to one.

Standard Normal Probability Distribution

To deal with problems where the normal probability distribution is applicable, the value of random variable x is standardized by expressing it as the number of standard deviations (σ) lying on both sides of its mean (μ). Such *standardized normal random variable*, z (also called *z–statistic*, *z–score* or *normal variate*) is defined as

$$z = \frac{x - \mu}{\sigma} \tag{7-8}$$

or equivalently $$x = \mu + z\sigma$$

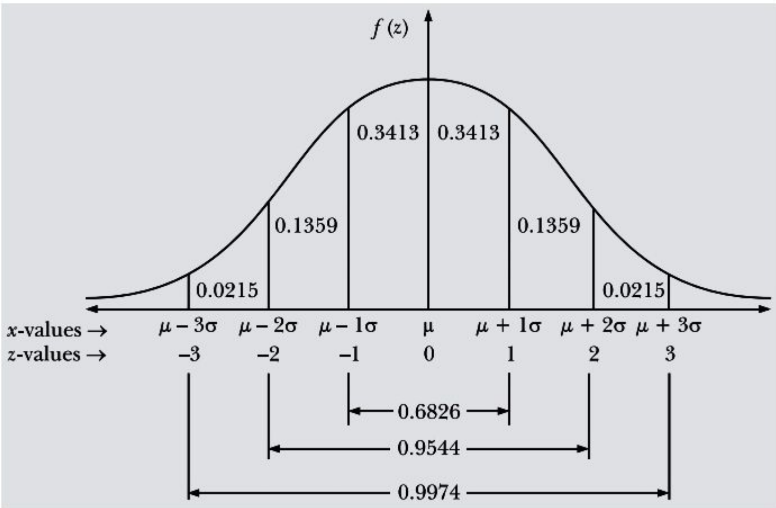
The z–statistic measures the number of standard deviations that any value of the random variable x falls from the mean. From formula (7.7), we may conclude that

(i) When x is less than the mean (μ), the value of z is negative.

(ii) When x is more than the mean (μ), the value of z is positive.

(iii) When $x = \mu$, the value of $z = 0$.

Figure 7.7
Standard Normal Distribution



Any normal probability distribution with parameters μ and σ can be converted into another distribution called **standard normal probability distribution** as shown in Fig. 7.7 with mean $\mu_z = 0$ and standard deviation $\sigma_z = 1$ with the help of the formula (7–7).

Since *z-statistic* measures the number of standard deviations that any value of the random variable x falls from the mean, therefore value of z (obtained by using the formula (7.7)) represents the area under the normal curve (see normal standard distribution in Appendix). For example, $z = \pm 2$ implies that the value of x is 2 standard deviations above or below the mean (μ).

Area Under the Normal Curve

Since tails of normal curve does not touch x-axis, the range of normal distribution is infinite in both the directions away from μ. This implies that as x moves away from μ, the pdf $f(x)$ approaches x-axis but never actually touches it.

The area under the standard normal distribution between the mean $z = 0$ and $z = z_0$ (a specified positive value of z) can be read from standard normal (z) table. For example, area between $1 \le z \le 2$ is the proportion of the area under the curve which lies between the vertical lines erected at two points along the x–axis. Also if the distance between x and μ is one standard deviation or $(x - \mu)/\sigma = 1$, then 34.134 per cent of the distribution lies between x and μ. Similarly, if x is at 2σ away from μ, i.e., $(x - \mu)/\sigma = 2$, then the area will include 47.725 per cent of the distribution and so on as shown in Table 7.9.

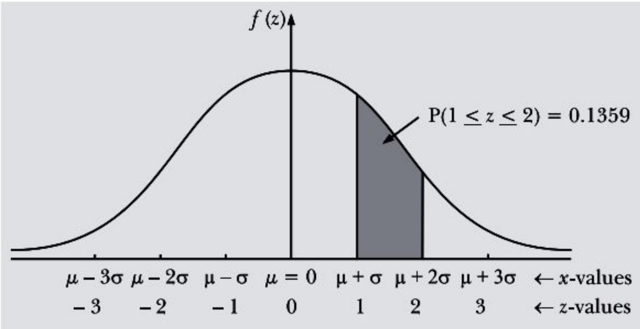


Figure 7.8
Diagram for Finding P(1 < z < 2)

Table 7.9 Area Under the Normal Curve

$z = \dfrac{x-\mu}{\sigma}$	*Area Under Normal Curve Between x and μ*
1.0	0.34134
2.0	0.47725
3.0	0.49875
4.0	0.49997

The percentage of the area of the normal distribution lying within the given range is shown in Table 7.10 and Fig. 7.7.

Table 7.10 Percentage of the Area of the Normal Distribution Lying within the Given Range

Number of Standard Deviations from Mean	*Approximate Percentage of Area under Normal Curve*
$x \pm \sigma$	68.26
$x \pm 2\sigma$	95.45
$x \pm 3\sigma$	99.75

Since standard normal distribution is a symmetrical distribution, therefore

$$P(0 \le z \le a) = P(-a \le z \le 0) \text{ for any value } a.$$

For example, $P(1 \le z \le 2) = P(z \le 2) - P(z \le 1) = 0.9772 - 0.8413 = 0.1359$

7.6.2 Approximation of Binomial and Poisson Distributions to Normal Distribution

The binomial distribution approaches a normal distribution with standardized variable, z, that is,

$$z = \frac{x - np}{\sqrt{npq}} \sim N(0, 1)$$

However, this approximation works well when both $np \ge 10$ and $npq \ge 10$

Similarly, Poisson distribution also approaches a normal distribution with standardized variable, z, that is,

$$z = \frac{x - \lambda}{\sqrt{\lambda}} \sim N(0, 1)$$

Example 7.27: How would you use the normal distribution to find approximate frequency of exactly 5 successes in 100 trials, the probability of success in each trial being $p = 0.1$.
Solution: Let n = number of trials, p = probability of success, and q = probability of failure.

Given $n = 100, p = 0.1$, and $q = 0.9$.

Hence, for binomial distribution, mean $\bar{x} = np = 100\times0.1 = 10$,

and standard deviation $(\sigma) = \sqrt{npq}$

$$= \sqrt{100 \times 0.1 \times 0.9} = 3.$$

When number of trials is large, binomial distribution tends to approximate normal distribution. Frequency of exactly 5 successes in 100 trials of binomial distribution will correspond to the frequency of class interval 4.5 to 5.5 and standard deviation of binomial distribution will correspond to mean and standard deviation of normal distribution;

$$z = -\frac{4.5 - np}{\sqrt{npq}} = \frac{4.5 - 10}{3} = -1.83.$$

Standard normal variate correspondent to 5.5 is

$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{5.5 - np}{\sqrt{npq}} = \frac{5.5 - 10}{3} = -1.50.$$

The area under normal curve between $z = -1.83$ and $z = -1.50$ is $0.668 - 0.336 = 0.0332$. Hence, approximate frequency of exactly 5 successes in 100 trials is $0.0332\times100 = 3.32$.

Example 7.28: When an aptitude test for selecting officers in a bank was conducted on 1,000 candidates, the average score is 42 and the standard deviation of scores is 24. Assuming normal distribution for the scores, find (a) number of candidates whose score exceeds 58, and (b) number of candidates whose scores lie between 30 and 66.

[*Karnataka Univ., B.Com., 2004*]

Solution: (a) Number of candidates whose score exceeds 58:

$$z = \frac{x - \mu}{\sigma} = \frac{58 - 42}{24} = 0.667.$$

Area under normal curve for $z = 0.667$ is $(0.5 - 0.2476) = 0.2524$. Thus the number of candidates whose score exceeds 58 is $1000 \times 0.2524 = 252.4$ or 252.

(b) Number of candidates whose scores lie between 30 and 66.
Standard normal variate corresponding to 30 is:

$$z_1 = \frac{x - \mu}{\sigma} = \frac{30 - 42}{24} = -0.5.$$

Standard normal variate corresponding to 66 is:

$$z_2 = \frac{x - \mu}{\sigma} = \frac{66 - 42}{24} = 1.$$

The area under normal curve between $z_1 = -0.5$ and $z_2 = 1$ is $1915+0.3413=0.5328$. Hence, the number of candidates whose scores lie between 30 and 66 is $1000\times0.5328 = 532.8$ or 533.

Example 7.29: There are 60 business students in the post-graduate department of a university, and the probability for any student to need a copy of particular textbook from the university library on any day is 0.05. How many copies of the book should be kept in the university library so that the probability may be greater than 0.90 that none of the students needing a copy from the library has to come back disappointed.

[*Kurukshetra Univ., M.Com., 2004*]

Solution: Let n be the number of students and p be the probability for any student to need a copy of a particular textbook from the university library. Then

Mean, $\mu = np = 600\times0.05 = 30$, and

$$\text{Standard deviation } (\sigma) = \sqrt{npq} = \sqrt{600 \times 0.05 \times 0.95} = 5.34.$$

Let x_1 be the number of copies of a textbook required on any day. Then

$$P(x_1) \geq 90\% = 0.9, \text{ i.e., } P\left[z = \frac{x_1 - \mu}{\sigma}\right] = P\left[z = \frac{x_1 - 30}{5.34}\right] \geq 1.28$$

$$x_1 - 30 \geq 6.835, \text{ i.e., } x_1 \geq 36.835 = 37.$$

[Since area under normal curve for $z_1 \geq 0.90 = 1.28$.]

Hence, the library should keep at least 37 copies of the book to ensure that the probability is more than 90 per cent that none of the students requiring a copy from the library has to come back disappointed.

Example 7.30: Of a large group of men, 4 per cent are under 60 inches in height and 40 per cent are between 60 and 45 inches. Assuming a normal distribution, find the mean height and standard deviation. *[Allahabad Univ., M.Com., 2006]*

Solution: Let x be the height of men, and μ be the the mean and σ be the standard deviation of the normal distribution. Then

$$P(x \leq 60) = P\left[z_1 \leq \frac{x-\mu}{\sigma} = \frac{60-\mu}{\sigma}\right] = 0.50 - 0.04 = 0.46.$$

The area under normal curve for $z_1 \leq 0.46$ is 1.645. Hence

$$P(x \leq 60) = P\left[z_1 \leq \frac{60-\mu}{\sigma}\right] = -1.645. \qquad (i)$$

The height of 40 per cent between 60 and 65 inches is equivalent to the area between $x = 60$ and $x = 65$ is 0.4 or the area to the left of the ordinate at $x = 65$. That is,

$$P(x \leq 65) = P\left[z_2 \leq \frac{65-\mu}{\sigma}\right] = 0.50 - 0.40 = 0.10.$$

The area under normal curve for $z_2 \leq 0.10$ is 0.13. Hence

$$P\left[z \leq \frac{65-\mu}{\sigma}\right] = -0.13 \qquad (ii)$$

Dividing Eqn. (*i*) by (*ii*), we get

$$\frac{60-\mu}{65-\mu} = \frac{1.645}{0.13} \text{ or } \mu = 65.42.$$

Putting the value of $\bar{x}$ in Eq. (*i*), we get $\sigma = \dfrac{60 - 12.65}{-1.645} = 3.29$. Hence $\bar{x} = 65.42$ and $\sigma = 3.29$.

Example 7.31: The customer accounts at a certain departmental store have an average balance of ₹480 and a standard deviation of ₹160. Assuming that the account balances are normally distributed.

(a) What proportion of the accounts is over ₹600?

(b) What proportion of the accounts is between ₹240 and ₹360? *[Madras Univ., M.Com., 2001]*

Solution: Let x be the balance of the customer accounts. For normal distribution, $\mu = 480$ and $\sigma = 160$. The standard normal variable z is

$$z = \frac{x-\mu}{\sigma} = \frac{x-480}{160}.$$

(a) If $x = 600$, then $z = \dfrac{600-480}{160} = 0.75$. Thus,

$$P(x > 600) = P(z > 0.75) = 0.5000 - 0.2734 = 0.2266.$$

Hence 22.66 per cent of the accounts have a balance in excees of ₹600.

(b) Probability that the accounts lie between ₹400 and ₹600 is given by $P(400 \leq x \leq 600)$.

If $x = 400$, then $z = \dfrac{400-480}{160} = -0.5$

If $x = 600$, then $z = \dfrac{600-480}{160} = -0.75$

So $P(400 < x < 600) = P(-0.5 < x < 0.75)$

$$= P(-0.5 \leq z \leq 0) + P(0 \leq z \leq 0.75) = 0.1915 + 0.2734 = 0.4649.$$

Hence 46.69 per cent of the accounts have an average balance between ₹400 and ₹600.

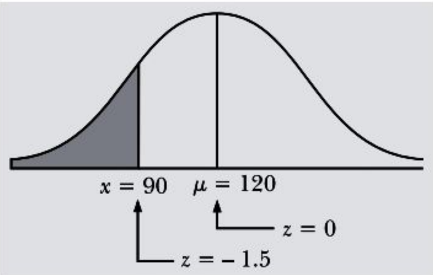
Example 7.32: 1000 light bulbs with a mean life of 120 days are installed in a new factory and their length of life is normally distributed with standard deviation of 20 days.

(a) How many bulbs will expire in less than 90 days?

(b) If it is decided to replace all the bulbs together, what interval should be allowed between replacements if not more than 10 per cent should expire before replacement?

Solution: (a) Given $\mu = 120$, $\sigma = 20$, and $x = 90$. Then

$$z = \frac{x-\mu}{\sigma} = \frac{90-120}{20} = -1.5$$



The area under the normal curve between $z = 0$ and $z = -1.5$ is 0.4332. Therefore, area to the left of -1.5 is $0.5 - 0.4332 = 0.0668$. Thus, the expected number of bulbs to expire in less than 90 days will be $0.0668 \times 1000 = 67$ (approx.).

(b) The value of z corresponding to an area 0.4 (0.5 – 0.10). Under the normal curve is 1.28. Therefore,

$$z = \frac{x-\mu}{\sigma} \quad \text{or} \quad -1.28 = \frac{x-120}{20} \quad \text{or} \quad x = 120 - 20(-1.28) = 94$$

Hence, the bulbs will have to be replaced after 94 days.

Example 7.33: The lifetimes of certain kinds of electronic devices have a mean of 300 hours and standard deviation of 25 hours. Assuming that the distribution of these lifetimes, which are measured to the nearest hour, can be approximated closely with a normal curve:

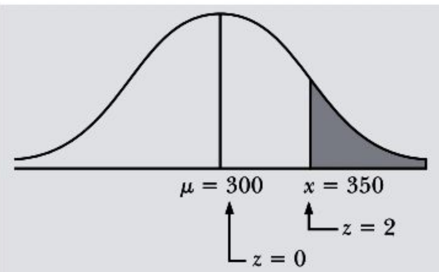
(a) Find the probability that any one of these electronic devices will have a lifetime of more than 350 hours.

(b) What percentage will have lifetimes of 300 hours or less?

(c) What percentage will have lifetimes from 220 or 260 hours?

Solution: (a) Given $\mu = 300$, $\sigma = 25$, and $x = 350$. Then

$$z = \frac{x-\mu}{\sigma} = \frac{350-300}{25} = 2$$



The area under the normal curve between $z = 0$ and $z = 2$ is 0.9772. Thus, the required probability is $1 - 0.9772 = 0.0228$

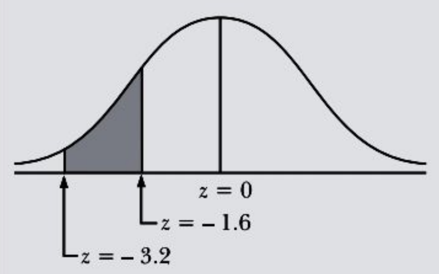
(b)
$$z = \frac{x-\mu}{\sigma} = \frac{300-300}{25} = 0$$

Therefore, the required percentage is $0.5000 \times 100 = 50$ per cent.

(c) Given $x_1 = 220$, $x_2 = 260$, $\mu = 300$ and $\sigma = 25$. Thus,

$$z_1 = \frac{220-300}{25} = -3.2$$

and
$$z_2 = \frac{260-300}{25} = -1.6$$



From the normal table, we have

$$P(z = -1.6) = 0.4452 \text{ and } P(z = -3.2) = 0.4903$$

Thus, the required probability is

$$P(z = -3.2) - P(z = -1.6) = 0.4903 - 0.4452 = 0.0541$$

Hence, the required percentage $= 0.0541 \times 100 = 5.41$ per cent.

Example 7.34: In a certain examination, the percentage of passes and distinctions were 46 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).

Also determine what would have been the minimum qualifying marks for admission to a re-examination of the failed candidates, had it been desired that the best 25 per cent of them should be given another opportunity of being examined.

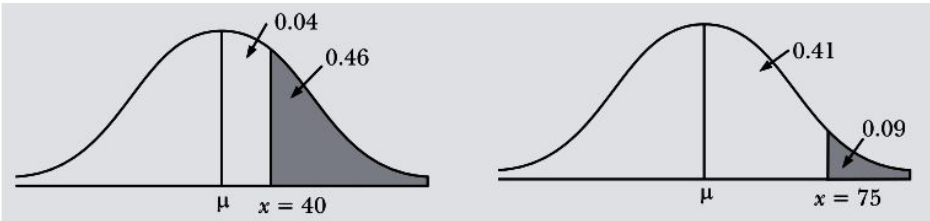
Solution: (a) Let μ be the mean and σ be the standard deviation of the normal distribution. The area to the right of the ordinate at $x = 40$ is 0.46 and hence the area between the mean and the ordinate at $x = 40$ is 0.04.

Now from the normal table, corresponding to 0.04, the standard normal variate, $z = 0.1$. Therefore, we have

$$\frac{40 - \mu}{\sigma} = 0.1 \text{ or } 40 - \mu = 0.1\sigma$$

Similarly, $$\frac{75 - \mu}{\sigma} = 1.34 \text{ or } 75 - \mu = 1.34\,\sigma$$

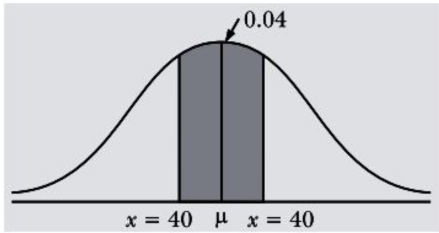
Solving these equations, we get $\sigma = 28.23$ and $\mu = 37.18$ or 37.



(b) Let us assume that x_1 is the minimum qualifying marks for re-examination of the failed candidates.

The area to the right of $x = 40$ is 46 per cent. Thus, the percentage of students failing is 54 and this is the area to the left of 40. We want that the best 25 per cent of these failed candidates should be given a chance to reappear. Suppose this area is equal to the shaded area in the diagram. This area is, 25 per cent of 54 = 13.5 per cent = 0.1350.

The area between mean and ordinate at $x_1 = -(0.1350 - 0.04) = -0.0950$ (negative sign is included because the area lies to the left of the mean ordinates).



Corresponding to this area, the standard normal variate $z = -0.0378$. Thus, we write

$$\frac{x_1 - \mu}{\sigma} = -0.0378$$

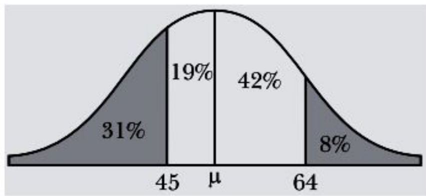
$$x_1 = \mu - 0.0378\,\sigma = 37.2 - (0.0378 \times 28.23)$$
$$= 37.2 - 1.067 = 36.133 \text{ or } 36 \text{ (approx.)}$$

Example 7.35: In a normal distribution 31 per cent of the items are under 45 and 8 per cent are over 64. Find the mean and standard deviation of the distribution.

[Delhi Univ., MBA, 2008]

Solution: Since 31 per cent of the items are under 45, therefore the left of the ordinate at $x = 45$ is 0.31, and obviously the area to the right of the ordinate up to the mean is $(0.5 - 0.31) = 0.19$. The value of z corresponding to this area is 0.5. Hence

$$z = \frac{45 - \mu}{\sigma} = -0.5 \text{ or } -\mu + 0.5\sigma = -45$$



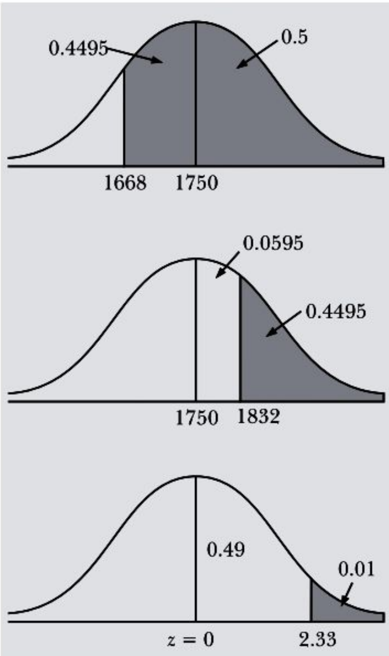
As 8 per cent of the items are above 64, therefore area to the right of the ordinate at 64 is 0.08. Area to the left of the ordinate at $x = 64$ up to mean ordinate is $(0.5 - 0.08) = 0.42$ and the value of z corresponding to this area is 1.4. Hence

$$z = \frac{64 - \mu}{\sigma} = 1.4 \text{ or } -\mu - 1.4\sigma = -64$$

From these two equations, we get $1.9\,\sigma = 19$ or $\sigma = 10$. Putting $\sigma = 10$ in the first equation, we get $\mu - 0.5 \times 10 = 45$ or $\mu = 50$. Thus, mean of the distribution is 50 and standard deviation is 10.

Example 7.36: The income of a group of 10,000 persons was found to be normally distributed with mean ₹1750 p.m. and standard deviation ₹50. Show that of this group 95 per cent had income exceeding ₹1668 and only 5 per cent had income exceeding ₹1832. What was the lowest income among the richest 100? *[Delhi Univ., MBA, 2007]*

Solution: (a) Given that $x = 1668$, $\mu = 1750$ and $\sigma = 50$. Therefore, the standard normal variate corresponding to $x = 1668$ is



$$z = \frac{x-\mu}{\sigma} = \frac{1668-1750}{50} = -1.64$$

The area to the right of the ordinate at $z = -1.64$ (or $x = 1668$) is (0.4495 + 0.5000) = 0.9495, because $z = -1.64$ to its right covers 95 per cent area).

The expected number of persons getting above ₹1668 are 10,000 × 0.9495 = 9495. This is about 95 per cent of the total of 10,000 persons.

(ii) The standard normal variate corresponding to $x = 1832$ is

$$z = \frac{1832 - 1750}{50} = 1.64$$

The area to the right of ordinate at $z = 1.64$ is 0.5000 – 0.4495 = 0.0505

The expected number of persons getting above ₹1832 is 10,000 × 0.0505 = 505. This is about 5 per cent of the total of 10,000 persons. Thus, probability of getting richest 100 out of 10,000 is 100/10,000 = 0.01.

The standard normal variate having 0.01 area to its right is $z = 2.33$. Hence,

$$2.33 = \frac{x - 1750}{50}$$

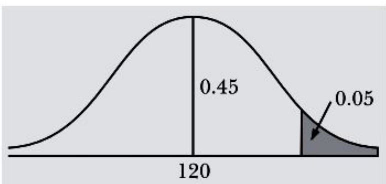
$$x = 2.33 \times 50 + 1750 = ₹1866 \text{ approx.}$$

This implies that the lowest among the richest 100 is getting ₹1866 per month.

Example 7.37: A wholesale distributor of fertilizer products finds that the annual demand for one type of fertilizer is normally distributed with a mean of 120 tonnes and standard deviation of 16 tonnes. If he orders only once a year, what quantity should be ordered to ensure that there is only a 5 per cent chance of running short? *[Delhi Univ., MBA, 2000, 2008]*

Solution: Let x be the annual demand (in tonnes) for one type of fertilizer. Therefore,

$$z = \frac{x - 120}{16}$$



The desired area of 5 per cent is shown in the figure. Since the area between the mean and the given value of x is 0.45, therefore from the normal table this area of 0.45 corresponds to $z = 1.64$.

Substituting this value of $z = 1.64$ in standard normal variate, we get

$$1.64 = \frac{x - 120}{16}$$

or $x = 120 + (1.64)(16) = 146.24$ tonnes.

If it is necessary to order in whole units, then the wholesale distributor should order 147 tonnes.

Example 7.38: Assume that the test scores from a college admissions test are normally distributed with a mean of 450 and a standard deviation of 100.

(a) What percentage of people taking the test score are between 400 and 500?

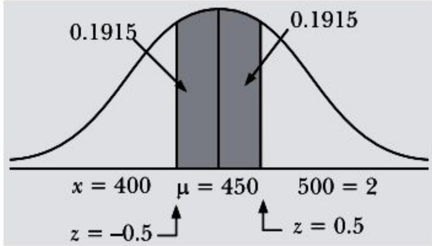
(b) Suppose someone received a score of 630. What percentage of the people taking the test score better? What percentage score worse?

(c) If a particular university will not admit any one scoring below 480, what percentage of the persons taking the test would be acceptable to the university?

[Delhi Univ., MBA, 2003]

Solution: (a) Given $\mu = 450$ and $\sigma = 100$. Let x be the test score. Then

$$z_1 = \frac{x-\mu}{\sigma} = \frac{500-450}{100} = 0.5$$



and $$z_2 = \frac{x-\mu}{\sigma} = \frac{630-450}{100} = -0.5$$

The area under the normal curve between $z = 0$ and $z = 0.5$ is 0.1915

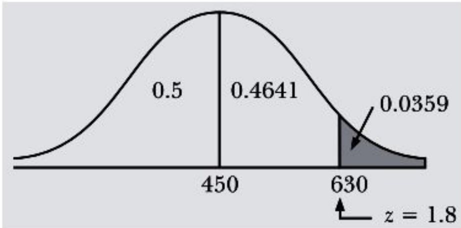
The required probability that the score falls between 400 and 500 is

$$P(400 \le x \le 500) = P(-0.5 \le z \le 0.5) = 0.1915 + 0.1915 = 0.3830$$

So, the percentage of the people taking the test score between 400 and 500 is 38.30 per cent.

(b) Given $x = 630$, $\mu = 450$ and $\sigma = 100$. Thus,

$$z = \frac{x-\mu}{\sigma} = \frac{630-450}{100} = 1.8$$



The area under the normal curve between $z = 0$ and $z = 1.8$ is 0.4641.

The probability that people taking the test score better is given by

$$P(x \ge 630) = P(z \ge 1.8) = 0.5000 + 0.4641 = 0.9640$$

That is, 96.40 per cent people score better.

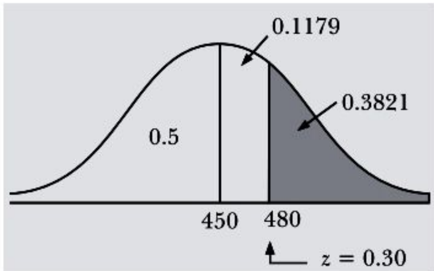
The probability that people taking the test score worse is given by

$$P(x \le 630) = P(z \le 1.8) = 0.5000 - 0.4641 = 0.0.0359$$

That is, 3.59 per cent people score worse.

(c) Given $x = 480$, $\mu = 450$ and $\sigma = 100$. Thus,

$$z = \frac{x-\mu}{\sigma} = \frac{480-450}{100} = 0.30$$



The area under the normal curve between $z=0$ and $z=0.30$ is 0.1179. So,

$$P(x \ge 480) = P(z \ge 0.30) = 0.5000 + 0.1179 = 0.6179$$

The percentage of people who score more than 480 and are acceptable to the university is 61.79 per cent.

Example 7.39: The results of particular examination are given below in a summary form:

Result	*Per cent of Candidates*
• Passed with distinction	10
• Passed with out distinction	60
• Failed	30

It is known that a candidate fails in the examination if he obtains less than 40 marks (out of 100) while he must obtain at least 75 marks in order to pass with distinction. Determine the mean and standard deviation of the distribution of marks, assuming this to be normal.

Solution: The given data are illustrated in the figure.

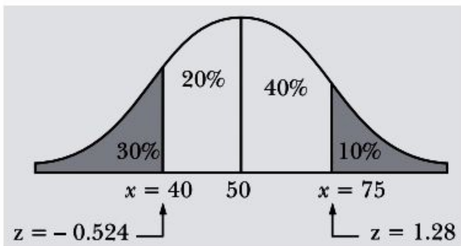
Since 30 per cent candidates who obtained less than 40 marks (out of 100) failed in the examination from the figure we have

$$z = \frac{x-\mu}{\sigma} \text{ or } -0.524 = \frac{40-\mu}{\sigma} \text{ or } \mu - 0.524\sigma = 40$$

(Table value of z corresponding to 20 per cent area under the normal curve is 0.524)

Also 10 per cent candidates who obtained more than 75 marks passed with distinction from the figure, we have

$$z = \frac{x-\mu}{\sigma} \text{ or } 1.28 = \frac{75-\mu}{\sigma} \text{ or } \mu + 1.28\sigma = 75$$



(Table value of z corresponding to 40 per cent area under normal curve is 1.28)

Solving these equations, we get mean $\mu = 50.17$ and standard deviation $\sigma = 19.4$

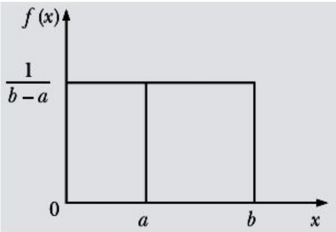
7.6.3 Uniform (Rectangular) Distribution

The simplest case of a continuous distribution is the uniform distribution. The general expression for the *pdf* (range of values) for a continuous random variable which is uniformly distributed over the interval between a to b is

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

This distribution is also known as *constant distribution* because the probability is constant $[= 1/(b - a)]$ at every point of the interval (a, b) and is independent of whatever value the variable may take within the interval.

Figure 7.9
Rectangular Probability Distribution



The general form of the rectangular probability distribution is shown in Fig. 7.9.

The mean and variance of this distribution are given by

$$\text{Mean} = \frac{(b+a)}{2} \text{ and Variance} = \frac{(b+a)^2}{12}$$

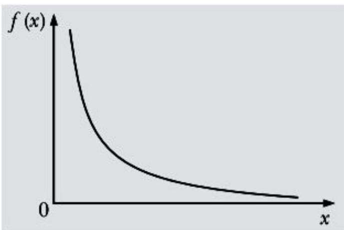
This distribution is useful when the probability of occurrences of an event is constant whatever be the value of the variable, i.e., all possible values of the continuous variable are assumed equally likely.

7.6.4 Exponential Probability Distribution

The probability density function (*pdf*) for exponential probability distribution is

$$f(x) = \begin{cases} \mu\, e^{-\mu x}, & x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

Figure 7.10
Exponential Probability Distribution



where μ (> 0) is a given parameter. This distribution is also referred to as *negative exponential distribution.* It is particularly useful in the queuing (waiting line) theory.

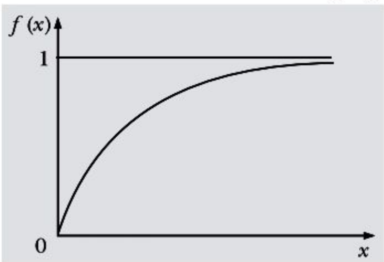
The graph of its *pdf* slopes downward to the right from its maximum at $x = 0$, where $f(x) = \mu$ as shown in Fig. 7.10.

The exponential distribution has the mean, $1/\mu$ and variance, $1/\mu^2$. The *cumulative density function* (*cdf*) of the exponential distribution is

$$F(x) = \int_0^x \mu\, e^{-\mu x}\, dx$$

$$= \Big[-e^{-\mu x}\Big]_0^x = 1 - e{-}\,\mu x$$

Figure 7.11
Exponential Probability Distribution



The graph of *cdf* is shown in Fig. 7.11.

The typical applications of *cdf* of exponential functions are found in representing a saturation phenomenon. That is, the situations where the effect of successive increments of the input x (e.g., size of advertising effort) show diminishing returns (e.g., resulting sales) as the total amount of x increases, and eventually, additional input increments have no effect.

Exponential distribution is closely related with the Poisson distribution. For example, if the Poisson random variable represents the *number of arrivals* per unit time at a service window, the exponential random variable will represent the *time between two successive arrivals*.

7.6.5 Gamma (Erlang) Distribution

The probability density function (*pdf*) for the gamma (or Erlang) distribution is

$$f(x) = \frac{\mu\,(\mu x)^{n-1} e^{-\mu x}}{(n-1)!},\ x > 0 \text{ and } \mu \ge 0$$

Gamma distribution is derived by the sum of n identically distributed and independent exponential random variables. Here, it may be noted that the *pdf* of gamma distribution reduces to the exponential density function for $n = 1$. This means the exponential distribution is the special case of the gamma distribution, where $n = 1$.

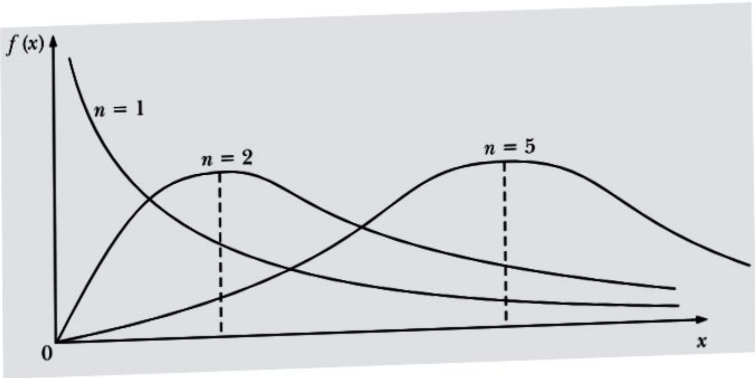


Figure 7.12 Gamma Distribution *pdf's* for $\mu = 1$

The graphs of the *pdf* 's for the gamma distribution for $\mu = 1$ and selected values of n are shown in Fig. 7.12.

In the gamma distribution *pdf* 's, the parameter μ changes the relative scales of the two axes, and the parameter n determines the location of the peak of the curve. However, for all values of these two parameters, the area under the curve is equal to 1.

The expected value and variance of this distribution are $E(x) = n/\mu$; $Var(x) = n/\mu^2$

7.6.6 Beta Distribution

The probability density function (*pdf*) for beta distribution is

$$f(x) = \frac{x^{m-1}(1-x)^{n-1}}{\beta(m, n)}; 0 \le x \le 1 ; m > 0; n > 0$$

where
$$\beta(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1}\, dx$$

is the beta function, whose value may be obtained directly from the table of the beta function.

The expected value and variance of random variable x in this case are given by

$$E(x) = \frac{m}{m+n} \text{ and } Var(x) = \frac{mn}{(m+n)^2(m+n+1)}$$

This distribution is commonly used to describe the random variable whose possible values lie in a restricted interval of numbers. A typical use of this distribution is found in the use of PERT where activity times are estimated within a specific range.

Conceptual Questions 7D

20. State the conditions under which a binomial distribution tends to (i) Poisson distribution, (ii) normal distribution. Write down the probability functions of binomial and Poisson distributions.

21. Normal distribution is symmetric with a single peak. Does this mean that all symmetric distributions are normal? Explain.

22. When finding probabilities with a normal curve we always deal with intervals; the probability of a single value of x is defined equal to zero. Why is this so?

23. When finding a normal probability, is there a difference between the values of $P(a < x < b)$ and $P(a \le x \le b)$, where a and b represent two numbers? Why or why not?

24. What are the parameters of normal distribution? What information is provided by these parameters?

25. What are the chief properties of normal distribution? Describe briefly the importance of normal distribution in statistical analysis. *[Delhi Univ., MBA, 2000]*

Formulae Used

1. Expected value of a random variable x

$$E(x) = \sum x.P(x)$$

where x = value of the random variable

$P(x)$ = probability that the random variable will take on the value x.

2. Binomial probability distribution

- Probability of r success in n Bernoulli trials

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

where p = probability of success

q = probability of failure, $q = 1 - p$

- Mean and standard deviation of binomial distribution

Mean $\mu = np$

Standard deviation $\sigma = \sqrt{npq}$

4. Poisson probability distribution

- Probability of getting exactly r occurrences of random event

$$P(x = r) = \frac{\lambda^r e^{-\lambda}}{r!}$$

where $\lambda = np$, mean number of occurrences per interval of time

e = 2.71828, a constant that represents the base of the natural logarithm system

- Mean and standard deviation of Poisson distribution

$$\lambda = np, \sigma = np$$

5. Normal distribution formula:

Number of standard deviations σ a value of random variable x is away from the mean μ of normal distribution:

$$z = \frac{x - \mu}{\sigma}$$

Chapter Concepts Quiz

True or False

1. [T] [F] The expected value of a random variable describes the long range weighted average of its values.
2. [T] [F] The mean of the binomial distribution is greater than its variance.
3. [T] [F] A binomial distribution is positively stewed when $p > 0.5$.
4. [T] [F] The mean, median, and mode always coincide in the normal distribution.
5. [T] [F] The expected value of a random variable is always a non-negative number.
6. [T] [F] The binomial distribution is symmetrical for any value of p (probability of success).
7. [T] [F] Poisson distribution generally describes arrivals at a service facility.
8. [T] [F] In a Bernoulli process, the probability of success must equal the probability of failure.
9. [T] [F] The symmetry of the normal distribution about its mean ensure that its tails extend indefinitely in both the positive and negative directions.
10. [T] [F] All normal distributions are defined by two measures – the mean and the standard deviation.
11. [T] [F] The expected value of a discrete random variable may be determined by taking an average of the values of the random variable.
12. [T] [F] Within 2σ limits from mean, the area under a normal curve is 95.45 per cent.
13. [T] [F] Any course of action that maximizes expected gain also minimizes expected loss.
14. [T] [F] The value of normal variate for some value of the random variable x lying in a normal distribution is the area between x and the mean μ of the distribution.
15. [T] [F] For a given binomial distribution with n fixed, if $p < 0.5$, then distribution will be skewed to the right.

Multiple Choice Questions

16. In a binomial distribution if n is fixed and $p > 0.5$, then
    (a) the distribution will be skewed to left
    (b) the distribution will be skewed to right
    (c) the distribution will be symmetric
    (d) cannot say anything
17. The binomial distribution is symmetric when:
    (a) $p < 0.5$ (b) $p > 0.5$
    (c) $p = 0.5$ (d) p has any value
18. Which of the following is the characteristic of the probability distribution of a random variable?
    (a) $0 \le P(A_i) \le 1$, for all i