UNIT 1 INTRODUCTION TO PROBABILITY

Structure

1.1 INTRODUCTION

In our daily lives, we face many situations when we are unable to forecast the future with complete certainty. That is, in many decisions, the uncertainty is faced. Need to cope up with the uncertainty leads to the study and use of the probability theory. The first attempt to give quantitative measure of probability was made by Galileo (1564-1642), an Italian mathematician, when he was answering the following question on the request of his patron, the Grand Duke of Tuscany, who wanted to improve his performance at the gambling tables: "With three dice a total of 9 and 10 can each be produced by six different combinations, and yet experience shows that the number 10 is oftener thrown than the number 9?" To the mind of his patron the cases were (1, 2, 6), (1, 3, 5), (1, 4, 4), (2, 2, 5), (2, 3, 4), (3, 3, 3) for 9 and (1, 3, 6), (1, 4, 5), (2, 2, 6), (2, 3, 5), (2, 4, 4), (3, 3, 4) for 10 and hence he was thinking that why they do not occur equally frequently i.e. why there chances are not the same? Galileo makes a careful analysis of all the cases which can occur, and he showed that out of the 216 possible cases 27 are favourable to the appearance of the number 10 since permutations of (1, 3, 6) are (1, 3, 6), (1, 6, 3), (3, 1, 6), (3, 6, 1), (6, 1, 3), (6, 3, 1) i.e. number of permutations of (1, 3, 6) is 6; similarly, the number of permutations of (1, 4, 5), (2, 2, 6), (2, 3, 5), (2, 4, 4), (3, 3, 4) is 6, 3, 6, 3, 3 respectively and hence the total number of cases come out to be 6 + 6 + 3 + 6 + 3 + 3 = 27 whereas the number of favourable cases for getting a total of 9 on three dice are
6 + 6 + 3 + 3 + 6 + 1 = 25. Hence, this was the reason for 10 appearing oftener thrown than 9. But the first foundation was laid by the two mathematicians Pascal (1623-62) and Fermat (1601-65) due to a gambler's dispute in 1654 which led to the creation of a mathematical theory of probability by them. Later, important contributions were made by various researchers including Huyghens (1629 - 1695), Jacob Bernoulli (1654-1705), Laplace (1749-1827), Abraham De-Moivre (1667-1754), and Markov (1856-1922). Thomas Bayes (died in 1761, at the age of 59) gave an important technical result known as Bayes' theorem, published after his death in 1763, using which probabilities can be revised on the basis of

some new information. Thereafter, the probability, an important branch of Statistics, is being used worldwide.

We will start this unit with very elementary ideas. In other words, we are assuming that reader knows nothing about probability. We will go step by step clearing the basic ideas which are required to understand the probability. In this unit, we will first present the various terms which are used in the definition of probability and then we will give the classical definition of probability and simple problems on it.

Objectives

After studing this unit, you should be able to:

- define and give examples of random experiment and trial;
- define and give examples of sample space, sample point and event;
- explain mutually exclusive, equally likely, exhaustive and favourable cases and why they are different in nature and how much these terms are important to define probability;
- explain the classical definition of probability;
- solve simple problems based on the classical definition of probability; and
- distinguish between odds in favour and odds against the happening of an event.

Random Experiment

An experiment in which all the possible outcomes are known in advance but we cannot predict as to which of them will occur when we perform the experiment, e.g. Experiment of tossing a coin is random experiment as the possible outcomes head and tail are known in advance but which one will turn up is not known.

Similarly, 'Throwing a die' and 'Drawing a card from a well shuffled pack of 52 playing cards 'are the examples of random experiment.

Trial

Performing an experiment is called trial, e.g.

(i) Tossing a coin is a trial.

(ii) Throwing a die is a trial.

1.2 SAMPLE SPACE, SAMPLE POINT AND EVENT

Sample Space

Set of all possible outcomes of a random experiment is known as sample space and is usually denoted by S, and the total number of elements in the sample space is known as size of the sample space and is denoted by n(S), e.g.

(i) If we toss a coin then the sample space is

S = {H, T}, where H and T denote head and tail respectively and n(S) = 2.

(ii) If a die is thrown, then the sample space is

$$S = \{1, 2, 3, 4, 5, 6\} \text{ and } n(S) = 6. \left[\begin{array}{l}\because \text{die has six faces}\\ \text{numbered } 1, 2, 3, 4, 5, 6\end{array}\right]$$

(iii) If a coin and a die are thrown simultaneously, then the sample space is
S = {H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6} and n(S) = 12.

where H1 denotes that the coin shows head and die shows 1 etc.

Note: Unless stated the coin means an unbiased coin (i.e. the coin which favours neither head nor tail).

(iv) If a coin is tossed twice or two coins are tossed simultaneously then the sample space is

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

where HH means both the coins show head, HT means the first coin shows head and the second shows tail, etc. Here, n(S) = 4.

(v) If a coin is tossed thrice or three coins are tossed simultaneously, then the sample space is

S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT} and n(S) = 8.

(vi) If a coin is tossed 4 times or four coins are tossed simultaneously then the sample space is

S = {HHHH, HHHT, HHTH, HTHH, THHH, HHTT, HTHT, HTTH, THHT, THTH, TTHH, HTTT, THTT, TTHT, TTTH, TTTT} and n(S) = 16.

(vii) If a die is thrown twice or a pair of dice is thrown simultaneously, then sample space is

S = {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1), (4, 2), (4, 3), (4, 4), (5, 5), (4, 6),
(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),
(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

here, e.g., (1, 4) means the first die shows 1 and the second die shows 4.

Here, n(S) = 36.

(viii) If a family contains two children then the sample space is

S = {B_1B_2, B_1G_2, G_1B_2, G_1G_2}

where B_i denotes that i^{th} birth is of boy, i = 1, 2, and

G_i denotes that i^{th} birth is of girl, i = 1, 2.

This sample space can also be written as

S = {BB, BG, GB, GG}

(ix) If a bag contains 3 red and 4 black balls and

(a) One ball is drawn from the bag, then the sample space is

$\{R_1, R_2, B_1, B_2, B_3, B_4\}$, where R_1, R_2, R_3 denote three red balls and B_1, B_2, B_3, B_4

denote four black balls in the bag.

(b) Two balls are drawn one by one without replacement from the bag, then the sample space is

$$S = \{R_1R_2, R_1R_3, R_1B_1, R_1B_2, R_1B_3, R_1B_4, R_2R_1, R_2R_3, R_2B_1, R_2B_2, R_2B_3, R_2B_4, R_3R_1, R_3R_2, R_3B_1, R_3B_2, R_3B_3, R_3B_4, B_1R_1, B_1R_2, B_1R_3, B_1B_2, B_1B_3, B_1B_4, B_2R_1, B_2R_2, B_2R_3, B_2B_1, B_2B_3, B_2B_4, B_3R_1, B_3R_2, B_3R_3, B_3B_1, B_3B_2, B_3B_4, B_4R_1, B_4R_2, B_4R_3, B_4B_1, B_4B_2, B_4B_3\}$$

Note: It is very simple to write the above sample space – first write all other balls with R_1, then with R_2, then with R_3 and so on.

Remark 1: If a random experiment with x possible outcomes is performed n times, then the total number of elements in the sample is x^n i.e. $n(S) = x^n$, e.g. if a coin is tossed twice, then $n(S) = 2^2 = 4$; if a die is thrown thrice, then $n(S) = 6^3 = 216$.

Now you can try the following exercise.

E1) Write the sample space if we draw a card from a pack of 52 playing cards.

Sample Point

Each outcome of an experiment is visualised as a sample point in the sample space. e.g.

(i) If a coin is tossed then getting head or tail is a sample point.

(ii) If a die is thrown twice, then getting (1, 1) or (1, 2) or (1, 3) or…or (6, 6) is a sample point.

Event

Set of one or more possible outcomes of an experiment constitutes what is known as event. Thus, an event can be defined as a subset of the sample space, e.g.

i) In a die throwing experiment, event of getting a number less than 5 is the set {1, 2, 3, 4},

which refers to the combination of 4 outcomes and is a sub-set of the sample space

= {1, 2, 3, 4, 5, 6}.

ii) If a card is drawn form a well-shuffled pack of playing cards, then the event of getting a card of a spade suit is

$\{1_S, 2_S, 3_S, ..., 9_S, 10_S, J_S, Q_S, K_S\}$

where suffix S under each character in the set denotes that the card is of spade and J, Q and K represent jack, queen and king respectively.

1.3 EXHAUSTIVE CASES, FAVOURABLE CASES, MUTUALLY EXCLUSIVE CASES AND EQUALLY LIKELY CASES

Exhaustive Cases

The total number of possible outcomes in a random experiment is called the exhaustive cases. In other words, the number of elements in the sample space is known as number of exhaustive cases, e.g.

(i) If we toss a coin, then the number of exhaustive cases is 2 and the sample space in this case is {H, T}.

(ii) If we throw a die then number of exhaustive cases is 6 and the sample space in this case is {1, 2, 3, 4, 5, 6}

Favourable Cases

The cases which favour to the happening of an event are called favourable cases. e.g.

(i) For the event of drawing a card of spade from a pack of 52 cards, the number of favourable cases is 13.
(ii) For the event of getting an even number in throwing a die, the number of favourable cases is 3 and the event in this case is {2, 4, 6}.

Mutually Exclusive Cases

Cases are said to be mutually exclusive if the happening of any one of them prevents the happening of all others in a single experiment, e.g.

(i) In a coin tossing experiment head and tail are mutually exclusive as there cannot be simultaneous occurrence of head and tail.

Equally Likely Cases

Cases are said to be equally likely if we do not have any reason to expect one in preference to others. If there is some reason to expect one in preference to others, then the cases will not be equally likely, For example,

(i) Head and tail are equally likely in an experiment of tossing an unbiased coin. This is because if someone is expecting say head, he/she does not have any reason as to why he/she is expecting it.

(ii) All the six faces in an experiment of throwing an unbiased die are equally likely.

You will become more familiar with the concept of “equally likely cases” from the following examples, where the non-equally likely cases have been taken into consideration:

(i) Cases of “passing” and “not passing” a candidate in a test are not equally likely. This is because a candidate has some reason(s) to expect “passing” or “not passing” the test. If he/she prepares well for the test, he/she will pass the test and if he/she does not prepare for the test, he/she will not pass. So, here the cases are not equally likely.

(ii) Cases of "falling a ceiling fan" and "not falling" are not equally likely. This is because, we can give some reason(s) for not falling if the bolts and other parts are in good condition.

1.5 CLASSICAL OR MATHEMATICAL PROBABILITY

Let there be 'n' exhaustive cases in a random experiment which are mutually exclusive as well as equally likely. Let 'm' out of them be favourable for the happening of an event A (say), then the probability of happening event A (denoted by P (A)) is defined as

$$P(A) = \frac{\text{Number of favourable cases for event A}}{\text{Number of exhaustive cases}} = \frac{m}{n} \qquad \ldots (1)$$

Probability of non-happening of the event A is denoted by $P\left(\overline{A}\right)$ and is defined as

$$P(\overline{A}) = \frac{\text{Number of favourable cases for event } \overline{A}}{\text{Number of exhaustive cases}} = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - P(A)$$

So, $\quad P\left(A\right) + P\left(\overline{A}\right) = 1$

Therefore, we conclude that, the sum of the probabilities of happening an event and that of its complementary event is 1.

Let us now prove that **$0 \le P(A) \le 1$**

Proof: We know that

$0 \le$ Number of favourable cases $\le$ No of exhaustive cases

[$\because$ Number of favourable cases can never be negative and can at the most be equal to the number of exhaustive cases.]

$\Rightarrow 0 \le m \le n$

Dividing both sides by n, we get

$$\Rightarrow \frac{0}{n} \le \frac{m}{n} \le \frac{n}{n}$$

$\Rightarrow 0 \le P(A) \le 1$

Remark 2: Probability of an impossible event is always zero and that of certain event is 1, e.g. probability of getting 7 when we throw a die is zero as getting 7 here is an impossible event and probability of getting either of the six faces is 1 as it is a certain event.

Classical definition of probability fails if

(i) The cases are not equally likely, e.g. probability of a candidate passing a test is not defined.

$$\left[\begin{array}{l}\text{Passing or failing in a test} \\ \text{are not equally likely cases.}\end{array}\right]$$

(ii) The number of exhaustive cases is indefinitely large, e.g.

probability of drawing an integer say 2 from the set of integers i.e.

$\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}$ by classical definition probability, is $\frac{1}{\infty}=0.$

But, in actual, it is not so, happening of 2 is not impossible, i.e. there are some chances of drawing 2. Hence, classical definition is failed here also.

Before we give some examples on classical definition of probability, let us take up some examples which define the events as subsets of sample space.

Example 1: If a fair die is thrown once, what is the event of?

(i) getting an even number

(ii) getting a prime number

(iii) getting a number multiple of 3

(iv) getting an odd prime

(v) getting an even prime

(vi) getting a number greater than 4

(vii) getting a number multiple of 2 and 3

Solution: When a die is thrown, then the sample space is

$$S = \{1, 2, 3, 4, 5, 6\}$$

(i) Let E_1 be the event of getting an even number,

$\therefore$ $E_1 = \{2, 4, 6\}$

(ii) Let E_2 be the event of getting a prime number,

$\therefore E_2 = \{2, 3, 5\}$

(iii) Let E_3 be the event of getting a number multiple of 3

$\therefore E_3 = \{3, 6\}$

(iv) Let E_4 be the event of getting an odd prime,

$\therefore E_4 = \{3, 5\}$

(v) Let E_5 be the event of getting an even prime,

$\therefore E_5 = \{2\}$

(vi) Let E_6 be the event of getting a number greater than 4,

$\therefore E_6 = \{5, 6\}$

(vii) Let E_7 be the event of getting a number multiple of 2 and 3,

$\therefore E_7 = \{6\}$

Example 2: If a pair of a fair dice is thrown, what is the event of

(i) getting a doublet

(ii) getting sum as 11

(iii) getting sum less than 5

(iv) getting sum greater than 16

(v) getting 3 on the first die

(vi) getting a number multiple of 3 on second die

(vii) getting a number multiple of 2 on first die and a multiple of 3 on second die.

Solution: When two dice are thrown, then the sample space is already given in (vii) of Sec.1.3.

(i) Let E_1 be the event of getting a doublet.

$\therefore E_1 = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

(ii) Let E_2 be the event of getting sum 11.

$\therefore E_2 = \{(5, 6), (6, 5)\}$

(iii) Let E_3 be the event of getting sum less than 5 i.e. sum can be 2 or 3 or 4

$\therefore E_3 = \{(1, 1), (1, 2), (2, 1), (1, 3), (3, 1), (2, 2)\}$

(iv) Let E_4 be the event of getting sum greater than 16.

$\therefore E_4 = \{\ \}$ i.e. E_4 is a null event.

(v) Let E_5 be the event of getting 3 on the first die i.e. 3 on first die and second die may have any number

$\therefore E_5 = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\}$

(vi) Let E_6 be the event of getting a number multiple of 3 on second die i.e. first die may have any number and the second has 3 or 6.

$$\therefore E_6 = \{(1, 3), (2, 3), (3, 3), (4, 3), (5, 3), (6, 3), (1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6)\}$$

(vii) Let E_7 be the event of getting a multiple of 2 on the first die and a multiple of 3 on the second die i.e. 2 or 4 or 6 on first die and 3 or 6 on the second.

$\therefore E_7 = \{(2, 3), (4, 3), (6, 3), (2, 6), (4, 6), (6, 6)\}$

Now, you can try the following exercise.

E2) If a die and a coin are tossed simultaneously, write the event of getting

(i) head and prime number

(ii) tail and an even number

(iii) head and multiple

1.6 SIMPLE PROBLEMS ON PROBABILITY

Now let us give some examples so that you become familiar as to how and when the classical definition of probability is used.

Example 3: A bag contains 4 red, 5 black and 2 green balls. One ball is drawn from the bag. Find the probability that?

(i) It is a red ball
(ii) It is not black
(iii) It is green or black

Solution: Let R_1, R_2, R_3, R_4 denote 4 red balls in the bag. Similarly B_1, B_2, B_3, B_4, B_5 denote 5 black balls and G_1, G_2 denote two green balls in the bag. Then the sample space for drawing a ball is given by

$$\{R_1, R_2, R_3, R_4, B_1, B_2, B_3, B_4, B_5, G_1, G_2\}$$

(i) Let A be the event of getting a red ball, then A = $\{R_1, R_2, R_3, R_4\}$

$$\therefore\ P(A) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{4}{11}$$

(ii) Let B be the event that drawn ball is not black, then B = $\{R_1, R_2, R_3, R_4, G_1, G_2\}$

$$\therefore\ P(B) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{6}{11}$$

(iii) Let C be the event that drawn ball is green or black, then C = $\{B_1, B_2, B_3, B_4, B_5, G_1, G_2\}$.

$$\therefore\ P(C) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{7}{11}$$

Example 4: Three unbiased coins are tossed simultaneously. Find the probability of getting

(i) at least two heads

(ii) at most two heads

(iii) all heads

(iv) exactly one head

(v) exactly one tail

Solution: The sample space in this case is

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

(i) Let E_1 be the event of getting at least 2 heads, then E_1 = {HHT, HTH, THH, HHH}

$$\therefore\ P(E_1) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{4}{8} = \frac{1}{2}$$

(ii) Let E_2 be the event of getting at most 2 heads then

E_2 = {TTT, TTH, THT, HTT, HHT, HTH, THH}

$$\therefore\ P(E_2) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{7}{8}$$

(iii) Let E_3 be the event of getting all heads, then

E_3 = {HHH}

$$\therefore\ P(E_3) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{1}{8}$$

(iv) Let E_4 be the event of getting exactly one head then

E_4 = {HTT, THT, TTH}

$$\therefore\ P(E_4) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{3}{8}$$

(v) Let E_5 be the event of getting exactly one tail, then

E_5 = {HHT, HTH, THH}

$$\therefore\ P(E_5) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{3}{8}$$

Example 5: A fair die is thrown. Find the probability of getting

(i) a prime number

(ii) an even number

(iii) a number multiple of 2 or 3

(iv) a number multiple of 2 and 3

(v) a number greater than 4

Solution: The sample space in this case is

S = {1, 2, 3, 4, 5, 6}

(i) Let E_1 be the event of getting a prime number, then

E_1 = {2, 3, 5}.

$$\therefore\ P(E_1) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{3}{6} = \frac{1}{2}$$

(ii) Let E_2 be the event of getting an even number, then

$E_2 = \{2, 4, 6\}$

$$\therefore \quad P(E_2) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{3}{6} = \frac{1}{2}$$

(iii) Let E_3 event of getting a multiple of 2 or 3, then

$E_3 = \{2, 3, 4, 6\}$

$$\therefore \quad P(E_3) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{4}{6} = \frac{2}{3}$$

(iv) Let E_4 event of getting a number multiple of 2 and 3, then

$E_4 = \{6\}$

$$\therefore \quad P(E_4) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{1}{6}$$

(v) Let E_5 be the event of getting a number greater than 4, then

$E_5 = \{5, 6\}$

$$\therefore \quad P(E_5) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{2}{6} = \frac{1}{3}$$

Example 6: In an experiment of throwing two fair dice, find the probability of getting

(i) a doublet
(ii) sum 7
(iii) sum greater than 8
(iv) 3 on first die and a multiple of 2 on second die
(v) prime number on the first die and odd prime on the second die.

Solution: The sample space has already been given in (vii) of Sec. 1.3.

Here, the sample space contains 36 elements i.e. number of exhaustive cases is 36.

(i) Let E_1 be the event of getting a doublet (i.e. same number on both dice), then

$E_1 = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$.

$$\therefore \quad P(E_1) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{6}{36} = \frac{1}{6}$$

(ii) Let E_2 be the event of getting sum 7, then

$E_2 = \{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\}$

$$\therefore \quad P(E_2) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{6}{36} = \frac{1}{6}$$

(iii) Let E_3 be the event of getting sum greater than 8, then

E_3 = {(3, 6), (6, 3), (4, 5), (5, 4), (4, 6),

(6, 4), (5, 5), (5, 6), (6, 5), (6, 6)}

$$\therefore \quad P(E_3) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{10}{36} = \frac{5}{18}$$

(iv) Let E_4 be the event of getting 3 on first die and multiple of 2 on second die, then

E_4 = {(3, 2), (3, 4), (3, 6)}

$$\therefore \quad P(E_4) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{3}{36} = \frac{1}{12}$$

(v) Let E_5 be the event of getting prime number on first die and odd prime on second die, then

E_5 = {(2, 3), (2, 5), (3, 3), (3, 5), (5, 3), (5, 5)}

$$\therefore \quad P(E_5) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{6}{36} = \frac{1}{6}$$

Example 7: Out of 52 well shuffled playing cards, one card is drawn at random. Find the probability of getting

(i) a red card

(ii) a face card

(iii) a card of spade

(iv) a card other than club

(v) a king

Solution: Here, the number of exhaustive cases is 52 and a pack of playing cards contains 13 cards of each suit (spade, club, diamond, heart).

(i) Let A be the event of getting a red card. We know that there are 26 red cards,

$$\therefore \quad P(A) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{26}{52} = \frac{1}{2}$$

(ii) Let B be the event of getting a face card. We know that there are 12 face cards (jack, queen and king in each suit),

$$\therefore \quad P(B) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{12}{52} = \frac{3}{13}$$

(iii) Let C be the event of getting a card of spade
We know that there are 13 cards of spade

$$\therefore\ P(C) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{13}{52} = \frac{1}{4}$$

(iv) Let D be the event of getting a card other than club.
As there are 39 cards other than that of club,.

$$\therefore\ P(D) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{39}{52} = \frac{3}{4}$$

(v) Let E be the event of getting a king.
We know that there are 4 kings,

$$\therefore\ P(E) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{4}{52} = \frac{1}{13}$$

Example 8: In a family, there are two children. Write the sample space and find the probability that

(i) the elder child is a girl

(ii) younger child is a girl

(iii) both are girls

(iv) both are of opposite sex

Solution: Let G_i denotes that i^{th} birth is of girl (i = 1, 2) and B_i denotes that i^{th} birth is of boy, (i =1, 2).

$$\therefore\ S = \{G_1G_2, G_1B_2, B_1G_2, B_1B_2\}$$

(i) Let A be the event that elder child is a girl

$$\therefore\ A = \{G_1G_2, G_1B_2\}$$

$$\text{and } P(A) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{2}{4} = \frac{1}{2}$$

(ii) Let B be the event that younger child is a girl

$$\therefore\ B = \{G_1G_2, B_1G_2\}$$

$$\text{and } P(B) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{2}{4} = \frac{1}{2}$$

(iii) Let C be the event that both the children are girls

$$\therefore\ C = \{G_1G_2\}$$

$$\text{and } P(C) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{1}{4}$$

(iv) Let D be the event that both children are of opposite sex

$\therefore$ D = $\{G_1B_2, B_1G_2\}$

and $P(D) = \dfrac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \dfrac{2}{4} = \dfrac{1}{2}$

Example 9: Find the probability of getting 53 sundays in a randomly selected non-leap year.

Solution: We know that there are 365 days in a non-leap year.

$\dfrac{365}{7} = 52\dfrac{1}{7}$ weeks

i.e. one non-leap year = (52 complete weeks + one over day). This over day may be one of the days

Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday

So, the number of exhaustive cases = 7

Let A be the event of getting 53 Sundays

There will be 53 Sundays in a non leap year if and only if the over day is Sunday.

$\therefore$ Number of favourable cases for event A= 1

$\therefore$ $P(A) = \dfrac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \dfrac{1}{7}$

Example 10: A single letter selected at random from the word 'STATISTICS'. What is the probability that it is a vowel?

Solution: Here, as the total number of letters in the word 'STATISTICS' is n = 10, and the number of vowels in the word is m = 3 (vowels are a, i, i),

$\therefore$ The required probability = $\dfrac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \dfrac{m}{n} = \dfrac{3}{10}$

Example 11: Three horses A, B and C are in a race. A is twice as likely to win as B and B is twice as likely to win as C. What are the respective probabilities of their winning?

Solution: Let p be the probability that A wins the race.

$\therefore$ Probability that B wins the race = twice the probability of A's winning

= 2(p) = 2 p

and probability of C's winning = twice the probability of B's winning

= 2(2p) = 4p

Now, as the sum of the probability of happening an event and that of its complementary event(s) is 1. Here, the complementary of A is the happening of B or C]

$\therefore$ p + 2p + 4p = 1, and hence $p = \dfrac{1}{7}$.

Therefore, the respective chances of winning A, B and C are $\frac{1}{7}, \frac{2}{7}$ and $\frac{4}{7}$.

Now, let us take up some problems on probability which are based on permutation/combination which you have already studied in Unit 4 of Course MST-001.

Example 12: Out of 52 well shuffled playing cards, two cards are drawn at random. Find the probability of getting.

(i) One red and one black

(ii) Both cards of the same suit

(iii) One jack and other king

(iv) One red and the other of club

Solution: Out of 52 playing cards, two cards can be drawn in ${}^{52}C_2$ ways i.e.

$\frac{52\times51}{2!} = 26\times51$ ways

(i) Let A be the event of getting one red and one black card, then the number of favourable cases for the event A are ${}^{26}C_1 \times {}^{26}C_1$ [As one red card out of 26 red cards can be drawn in ${}^{26}C_1$ ways and one black card out of 26 black cards can be drawn in ${}^{26}C_1$ ways.]

$$\therefore \ P(A) = \frac{{}^{26}C_1 \times {}^{26}C_1}{{}^{52}C_2} = \frac{26\times26}{26\times51} = \frac{26}{51}$$

(ii) Let B be the event of getting both the cards of the same suit and i.e. two cards of spade or two cards of club or 2 cards of diamond or 2 cards of heart.

$\therefore$ Number of favourable cases for event B $= {}^{13}C_2 + {}^{13}C_2 + {}^{13}C_2 + {}^{13}C_2$

$$= 4\times{}^{13}C_2 = 4\times\frac{13\times12}{2!} = 2\times13\times12$$

$$\therefore \ P(B) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{2\times13\times12}{26\times51} = \frac{4}{17}$$

(iii) Let C be the event of getting a jack and a king.

$$\therefore P(C) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{{}^{4}C_1\times{}^{4}C_1}{{}^{52}C_2} = \frac{4\times4}{26\times51} = \frac{8}{663}$$

(iv) Let D be the event of getting one red and one card of club.

$$\therefore P(D) = \frac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \frac{{}^{26}C_1\times{}^{13}C_1}{{}^{52}C_2} = \frac{26\times13}{26\times51} = \frac{13}{51}$$

Example 13: If the letters of the word STATISTICS are arranged randomly then find the probability that all the three T's are together.

Solution: Let E be the event that selected word contains 3 T's together.

There are 10 letters in the word STATISTICS. If we consider three T's as a single letter $\boxed{\text{TTT}}$, then we have 8 letters i.e.1 '$\boxed{\text{TTT}}$'; 3 'S' ; 1 'A'; 2 'I' and 1 'C'

Number of possible arrangements with three T's coming together $= \dfrac{8!}{2!.3!}$

Number of favourable cases for event E $= \dfrac{8!}{2!.3!}$ and

Number of exhaustive cases = Total number of permutations of 10 letters in the word STATISTICS

$$= \frac{10!}{2!.3!.3!}$$

$\left[\because \text{out of 10 letters, 3 are T's, 2 are I's and 3 are S's}\right]$

$$P(A) = \frac{\dfrac{8!}{2!.3!}}{\dfrac{10!}{2!.3!.3!}} = \frac{8!.3!}{10!} = \frac{8!\times 6}{10\times 9\times 8!} = \frac{6}{10\times 9} = \frac{1}{15}$$

Example 14: In a lottery, one has to choose six numbers at random out of the numbers from 1 to 30. He/ she will get the prize only if all the six chosen numbers matched with the six numbers already decided by the lottery committee. Find the probability of wining the prize.

Solution: Out of 30 numbers 6 can be drawn in

$${}^{30}C_6 = \frac{30\times 29\times 28\times 27\times 26\times 25}{6!} = \frac{30\times 29\times 28\times 27\times 26\times 25}{720} = 593775 \text{ ways}$$

$\therefore$ Number of exhaustive cases = 593775

Out of these 593775 ways, there is only one way to win the prize (i.e. choose those six numbers that are already fixed by committee).

Here, the number of favourable cases is 1.

Hence, $P(\text{wining the prize}) = \dfrac{\text{Favourable cases}}{\text{Exhaustive cases}} = \dfrac{1}{593775}$

Now, you can try the following exercises.

E3) If two coins are tossed then find the probability of getting.

(i) At least one head

(ii) head and tail

(iii) At most one head

E4) If three dice are thrown, then find the probability of getting

(i) triplet
(ii) sum 5
(iii) sum at least 17
(iv) prime number on first die and odd prime number on second and third dice.

E5) Find the probability of getting 53 Mondays in a randomly selected leap year.

1.7 CONCEPT OF ODDS IN FAVOUR OF AND AGAINST THE HAPPENING OF AN EVENT

Let n be the number of exhaustive cases in a random experiment which are mutually exclusive and equally likely as well. Let m out of these n cases are favourable to the happening of an event A (say). Thus, the number of cases against A are $n-m$

Then odds in favour of event A are $m:n-m$ (i.e. m ratio $n-m$) and odds against A are $n-m:m$ (i.e. $n-m$ ratio m)

Example 15: If odds in favour of event A are $3:4$, what is the probability of happening A?

Solution: As odds in favour of A are $3:4$,

$\therefore$ $m=3$ and $n-m=4$ implies that $n=7$. Thus,

Probability of happening A i.e. $P(A) = \frac{m}{n} = \frac{3}{7}$.

Example 16: Find the probability of event A if

(i) Odds in favour of event A are 4 : 3

(ii) Odds against event A are 5 : 8

Solution: (i) We know that if odds in favour of A are $m:n$, then

$$P(A)=\frac{m}{m+n} \Rightarrow P(A) = \frac{4}{4+3} = \frac{4}{7}$$

(ii) Here, $n-m=5$ and $m=8$, therefore, $n=5+8=13$.

Now, as we know that if odds against the happening of an event A are $n-m:n$, then

$$P(A)=\frac{m}{n} \Rightarrow P(A) = \frac{8}{13}$$

Example 17 If $P(A) = \frac{3}{5}$ then find

(i) odds in favour of A; (ii) odds against the happening of event A.

Solution: (i) As $P(A) = \frac{3}{5}$,

$\therefore$ odds in favour of A in this case are $3:5-3 = 3:2$

(ii) We know that if P(A) = $\frac{m}{n}$, then odds against the happening of A are $n-m:m$

$\therefore$ In this case odds against the happening of event A are $5-3 : 3 = 2 : 3$

Now, you can try the following exercises.

E6) The odds that a person speaks the truth are 3 : 2. What is the probability that the person speaks truth?

E7) The odds against Manager X setting the wage dispute with the workers are 8 : 6. What is the probability that the manager settles the dispute?

E8) The probability that a student passes a test is $\frac{2}{3}$. What are the odds against passing the test by the student?

E9) Find the probability of the event A if

(i) Odds in favour of the event $\overline{A}$ are $1:4$ (ii) Odds against the event $\overline{A}$ are $7:2$

1.8 SUMMARY

Let us now summarize the main points which have been covered in this unit.

1) An experiment in which all the possible outcomes are known in advance but we cannot predict as to which of them will occur when we perform the experiment is called **random experiment**. Performing an experiment is called **trial**.
2) Set of all possible outcomes of a random experiment is known as **sample space**. Each outcome of an experiment is visualised as a **sample point** and set of one or more possible outcomes constitutes what is known as **event**. The total number of elements in the sample space is called the number of **exhaustive cases** and number of elements in favour of the event is the number of **favourable cases** for the event.
3) Cases are said to be **mutually exclusive** if the happening of any one of them prevents the happening of all others in a single experiment and if we do not have any reason to expect one in preference to others, then they are said to be **equally likely**.
4) **Classical Probability** of happening of an event is the ratio of number of favourable cases to the number of exhaustive cases, provided they are equally likely, mutually exclusive and finite.
5) **Odds in favour of an event** are the number of favourable cases: number of cases against the event, whereas **Odds against the event** are the number of cases against the event : number of cases favourable to the event.

1.9 SOLUTIONS/ANSWERS

E 1) Let suffices C, D, S, H denote that corresponding card is a club, diamond, spade, heart respectively then sample space of drawing a card can be written as

$$\{1_C, 2_C, 3_C, ..., 9_C, 10_C, J_C, Q_C, K_C, 1_D, 2_D, 3_D, ..., 9_D, 10_D, J_D, Q_D, K_D,$$
$$1_S, 2_S, 3_S, ..., 9_S, 10_S, J_S, Q_S, K_S, 1_H, 2_H, 3_H, ..., 9_H, 10_H, J_H, Q_H, K_H\}$$

E 2) If a die and a coin are tossed simultaneously then sample space is

{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6}

(i) Let A be the event of getting head and prime number, then A = {H2, H3, H5}

(ii) Let B be the event of getting tail and even number, then B = {T2, T4, T6}

(iii) Let C be the event of getting head and multiple of 3, then C = {H3, H6}

E 3) When two coins are tossed simultaneously then sample space is

S = {HH, HT, TH, TT}

(i) Let A be the event of getting at least one head, then

A = {HH, HT, TH} and $P(A) = \frac{3}{4}$

(ii) Let B be the event of getting both head and tail, then

B = {HT, TH} and $P(B) = \frac{2}{4} = \frac{1}{2}$

(iii) Let C be the event of getting at most one head, then

C = {TT, TH, HT} and $P(C) = \frac{3}{4}$

E 4) When 3 dice are thrown, then the sample space is

S = {(1, 1, 1), (1, 1, 2), (1, 1, 3), … , (1, 1, 6),
(1, 2, 1), (1, 2, 2), (1, 2, 3), … , (1, 2, 6),
(1, 3, 1), (1, 3, 2), (1, 3, 3), … , (1, 3, 6),
.
.
.
(6, 6, 1), (6, 6, 2), (6, 6, 3), …, (6, 6, 6)}

Number of elements in the sample space = $6 \times 6 \times 6 = 216$

$$\left[\begin{array}{l}\because \text{We have to fill up 3 positions } (.,.,.) \text{ and each position can be} \\ \text{filled with 6 options, this can be done in } 6\times6\times6 = 216 \text{ ways}\end{array}\right]$$

(i) Let A be the event of getting triplet i.e. same number on each die.

$$\therefore A = \{(1,1,1), (2,2,2), (3,3,3), (4,4,4), (5,5,5), (6,6,6)\}$$

and hence P(A) = $\dfrac{6}{216} = \dfrac{1}{36}$

(ii) Let B be the event of getting sum 5

$$\therefore B = \{(1,1,3), (1,3,1), (3,1,1), (1,2,2), (2,1,2), (2,2,1)\}$$

and P (B) = $\dfrac{6}{216} = \dfrac{1}{36}$

(iii) Let C be the event of getting sum at least17 i.e. sum 17 or 18

$$\therefore C = \{(5,6,6), (6,5,6), 6,6,5), (6,6,6)\}$$

and hence P (C) = $\dfrac{4}{216} = \dfrac{1}{54}$

(iv) Let D be the event of getting prime number on first die and odd prime number on second and third dice.
i.e. first die can show 2 or 3 or 5 and second, third dice can show 3 or 5

$$\therefore D = \{(2,3,3), (2,3,5), (2,5,3), (2,5,5), (3,3,3), (3,3,5), \\ (3,5,3), (3,5,5), (5,3,3), (5,3,5), (5,5,3), (5,5,5)\}$$

and hence P (D) = $\dfrac{12}{216} = \dfrac{1}{18}$

E5) We know that there are 366 days in a leap year. i.e. $\dfrac{365}{7} = 52\dfrac{2}{7}$ weeks

i.e one leap year = (52 complete weeks + two over days).

These two over days may be

(i) Sunday and Monday

(ii) Monday and Tuesday

(iii) Tuesday and Wednesday

(iv) Wednesday and Thursday

(v) Thursday and Friday

(vi) Friday and Saturday

(vii) Saturday and Sunday

$\therefore$ Number of exhaustive Cases = 7

Let A be the event of getting 53 Mondays

There will be 53 Mondays in a leap year if and only if these two over days are

"Sunday and Monday" or "Monday and Tuesday"

$\therefore$ Number of favourable cases for event A are 2

and $P(A) = \dfrac{\text{Number of favourable cases}}{\text{Number of exhaustive cases}} = \dfrac{2}{7}$

E6) Here, the odds in favour of speaking the truth are 3 : 2,

$\therefore$ here m = 3, n – m = 2 and hence n = 5.

Hence, the probability of speaking the truth $= \dfrac{3}{5}$

E7) As the odds against Manager X setting the wage dispute with the workers are 8 : 6, hence odds in favour of settling the dispute are 6 : 8.

Thus, the probability that the manager settles the dispute $= \dfrac{6}{14} = \dfrac{3}{7}$.

E8) As the probability that student pass a test = 2/3,

$\therefore$ the number of favourable cases = 2 and the number of exhaustive cases = 3, and hence the number of cases against passing the test = 3 – 2 = 1.

Thus, odds against passing the test

= the number cases against the event : the number cases favourable to the event

= 1 : 2

E 9) (i) Odds in favour of the event $\overline{A}$ are given as 1 : 4.

We know that if odds in favour of event E are $m:n$ then P(E) $= \dfrac{m}{m+n}$

$\therefore$ In this case $P(\overline{A}) = \dfrac{1}{1+4} = \dfrac{1}{5}$ and $P(A) = 1 - P(\overline{A}) = 1 - \dfrac{1}{5} = \dfrac{4}{5}$

(ii) Odds against the happening of the event $\overline{A}$ are given as 7 : 2.

We know that if odds against the happening of an event E are $m : n$, then $P(E) = \dfrac{n}{m+n}$.

$\therefore$ In this case $P(\overline{A}) = \dfrac{2}{7+2} = \dfrac{2}{9}$ and hence

$P(A) = 1 - P(\overline{A}) = 1 - \dfrac{2}{9} = \dfrac{7}{9}$.

UNIT 2 DIFFERENT APPROACHES TO PROBABILITY THEORY

Structure

2.1 INTRODUCTION

In the previous unit, we have defined the classical probability. There are some restrictions in order to use it such as the outcomes must be equally likely and finite. There are many situations where such conditions are not satisfied and hence classical definition cannot be applied. In such a situation, we need some other approaches to compute the probabilities.

Thus, in this unit, we will discuss different approaches to evaluate the probability of a given situation based on past experience or own experience or based on observed data. Actually classical definition is based on the theoretical assumptions and in this unit, our approach to evaluate the probability of an event is different from theoretical assumptions and will put you in a position to answer those questions related to probability where classical definition does not work. The unit discusses the relative frequency (statistical or empirical probability) and the subjective approaches to probability. These approaches, however, share the same basic axioms which provide us with the unified approach to probability known as axiomatic approach. So, the axiomatic approach will also be discussed in the unit.

Objectives

After studying this unit, you should be able to:

- explain the relative frequency approach and statistical(or empirical) probability;
- discuss subjective approach to probability; and
- discuss axiomatic approach to probability.

2.2 RELATIVE FREQUENCY APPROACH AND STATISTICAL PROBABILITY

Classical definition of probability fails if

i) the possible outcomes of the random experiment are not equally likely or/and

ii) the number of exhaustive cases is infinite.

In such cases, we obtain the probability by observing the data. This approach to probability is called the relative frequency approach and it defines the statistical probability. Before defining the statistical probability, let us consider the following example:

Following table gives a distribution of daily salary of some employees:

Salary per day (In Rs)	**Below 100**	**100-150**	**150-200**	**200 and above**
Employees	**20**	**40**	**50**	**15**

If an individual is selected at random from the above group of employees and we are interested in finding the probability that his/her salary was under Rs. 150, then as the number of employees having salary less than Rs 150 is 20 + 40 = 60 and the total number employees is 20 + 40 + 50 + 15 = 125, therefore the relative frequency that the employee gets salary less than Rs. 150 is

$$\frac{60}{125} = \frac{12}{25}.$$

This relative frequency is nothing but the probability that the individual selected is getting the salary less than Rs. 150.

So, in general, if X is a variable having the values $x_1, x_2, \ldots, x_n$ with frequencies $f_1, f_2, \ldots, f_n,$ respectively. Then

$$\frac{f_1}{\sum_{i=1}^{n} f_i}, \frac{f_2}{\sum_{i=1}^{n} f_i}, \ldots, \frac{f_n}{\sum_{i=1}^{n} f_i}$$

are the relative frequencies of $x_1, x_2, \ldots, x_n$ respectively and hence the probabilities of X taking the values $x_1, x, \ldots, x_n$ respectively.

But, in the above example the probability has been obtained using the similar concept as that of classical probability.

Now, let us consider a situation where a person is administered a sleeping pill and we are interested in finding the probability that the pill puts the person to sleep in 20 minutes. Here, we cannot say that the pill will be equally effective for all persons and hence we cannot apply classical definition here.

To find the required probability in this case, we should either have the past data or in the absence of the past data, we have to undertake an experiment where we administer the pill on a group of persons to check the effect. Let m be the number of persons to whom the pill put to sleep in 20 minutes and n be the total number of persons who were administered the pill.

Then, the relative frequency and hence the probability that a particular person will put to sleep in 20 minutes is $\frac{m}{n}$. But, this measure will serve as probability only if the total number of trials in the experiment is very large.

In the relative frequency approach, as the probability is obtained by repetitive empirical observations, it is known as statistical or empirical probability.

Statistical (or Empirical) Probability

If an event A (say) happens m times in n trials of an experiment which is performed repeatedly under essentially homogeneous and identical conditions (e.g. if we perform an experiment of tossing a coin in a room, then it must be performed in the same room and all other conditions for tossing the coin should also be identical and homogeneous in all the tosses), then the probability of happening A is defined as:

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

As an illustration, we tossed a coin 200 times and observed the number of heads. After each toss, proportion of heads i.e. $\frac{m}{n}$ was obtained, where m is the number of heads and n is the number of tosses as shown in the following table (Table 2.1):

Table 2.1: Table Showing Number of Tosses and Proportion of Heads

n (Number of Tosses)	m (Number of Heads)	Proportion of Heads i.e.P(H)=m/n
1	1	1
2	2	1
3	2	0.666667
4	3	0.75
5	4	0.8
6	4	0.666667
7	4	0.571429
8	5	0.625
9	6	0.666667
10	6	0.6
15	10	0.666667
20	12	0.6
25	14	0.56
30	16	0.533333
35	18	0.514286

40	22	0.55
45	25	0.555556
50	29	0.58
60	33	0.55
70	41	0.585714
80	46	0.575
90	52	0.577778
100	53	0.53
120	66	0.55
140	72	0.514286
160	82	0.5125
180	92	0.511111
200	105	0.525

Then a graph was plotted taking number of tosses (n) on x-axis and proportion of heads $\left(\frac{m}{n}\right)$ on y-axis as shown in Fig. 2.1.

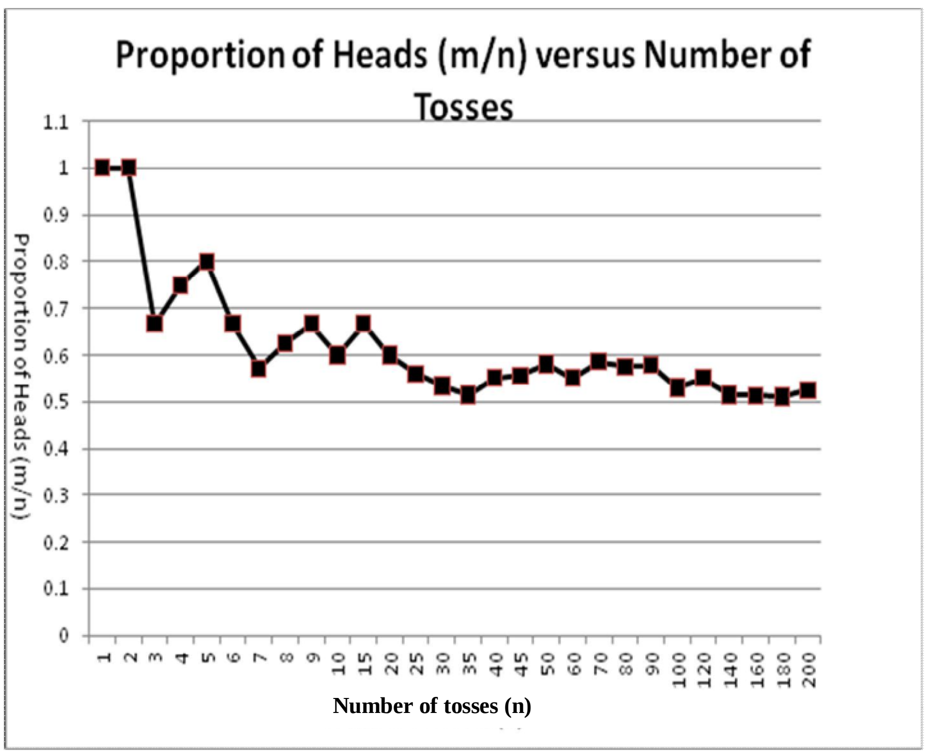


Fig. 2.1: **Proportion of Heads versus Number of Tosses**

The Graph reveals that as we go on increasing n,

$\frac{m}{n}$ tends to $\frac{1}{2}$

i.e. $\lim_{n\to\infty} \frac{m}{n} = \frac{1}{2}$

Hence, by the statistical (or empirical) definition of probability, the probability of getting head is

$$\lim_{n\to\infty} \frac{m}{n} = \frac{1}{2}.$$

Statistical probability has the following limitations:

(i) The experimental condition may get altered if it is repeated a large number of times.

(ii) $\lim_{n\to\infty} \frac{m}{n}$ may not have a unique value, however large n may be.

2.3 PROBLEMS BASED ON RELATIVE FREQUENCY

Example 1: The following data relate to 100 couples

Age of wife / Age of Husband	10-20	20-30	30-40	40-50	50-60
15-25	6	3	0	0	0
25-35	3	16	10	0	0
35-45	0	10	15	7	0
45-55	0	0	7	10	4
55-65	0	0	0	4	5

(i) Find the probability of a couple selected at random has a "age of wife" in the interval 20-50.

(ii) What is the probability that the age of wife is in the interval 20-40 and the age of husband is in the interval 35-45 if a couple selected at random?

Solution: (i) Required probability is given by

$$= \frac{(3+16+10+0+0)+(0+10+15+7+0)+(0+0+7+10+4)}{100}$$

$$= \frac{82}{100} = 0.82$$

(ii) Required probabilty $= \frac{10+15}{100} = \frac{25}{100} = 0.25$

Example 2: A class has 15 students whose ages are 14, 17, 15, 21, 19, 20, 16, 18, 20, 17, 14, 17, 16, 19 and 20 years respectively. One student is chosen at random and the age of the selected student is recorded. What is the probability that

(i) the age of the selected student is divisible by 3,

(ii) the age of the selected student is more than 16, and

(iii) the selected student is eligible to pole the vote.

Solution:

Age X	Frequency f	Relative frequency
14	2	2/15
15	1	1/15
16	2	2/15
17	3	3/15
18	1	1/15
19	2	2/15
20	3	3/15
21	1	1/15

(i) The age divisible by 3 is 15 or 18 or 21.

$$\therefore \text{ Required Probability} = \frac{1+1+1}{15} = \frac{3}{15} = \frac{1}{5}$$

(ii) Age more than 16 means, age may be 17, 18, 19, 20, 21

$$\therefore \text{ Required Probability} = \frac{3+1+2+3+1}{15} = \frac{10}{15} = \frac{2}{3}$$

(iii) In order to poll the vote, age must be ≥ 18 years. Thus, we are to obtain the probability that the selected student has age 18 or 19 or 20 or 21.

$$\therefore \text{ Required Probability} = \frac{1+2+3+1}{15} = \frac{7}{15}$$

Example 3: A tyre manufacturing company kept a record of the distance covered before a tyre needed to be replaced. The following table shows the results of 2000 cases.

Distance (in km)	Less than 4000	4001-10000	100001-20000	20001-40000	More than 40000
Frequency	20	100	200	1500	180

If a person buys a tyre of this company then find the probability that before the need of its replacement, it has covered

(i) at least a distance of 4001 km.

(ii) at most a distance 20000 km

(iii) more than a distance 20000 km

(iv) a distance between 10000 to 40000

Solution: The record is based on 2000 cases,

$\therefore$ Exhaustive cases in each case = 2000

(i) Out of 2000 cases, the number of cases in which tyre covered at least 4001 km

= 100 + 200 + 1500 + 180 = 1980

$$\therefore \text{Required Probability} = \frac{1980}{2000} = \frac{198}{200} = \frac{99}{100}$$

(ii) Number of cases in which distance covered by tyres of this company is

at most 20000 km = 20 + 100 + 200 = 320

$$\therefore \text{Required Probability} = \frac{320}{2000} = \frac{32}{200} = \frac{4}{25}$$

(iii) Number of cases in which tyres of this company covers a distance

of more than 20000 = 1500 + 180 = 1680

$$\therefore \text{Required Probability} = \frac{1680}{2000} = \frac{168}{200} = \frac{21}{25}$$

(iv) Number of cases in which tyres of this company covered a distance

between 10000 to 40000 = 200 + 1500 = 1700

$$\therefore \text{Required Probability} = \frac{1700}{2000} = \frac{17}{20}$$

Now, you can try the following exercises.

E 1) An insurance company selected 5000 drivers from a city at random in order to find a relationship between age and accidents. The following table shows the results related to these 5000 drivers.

Age of driver (in years)	Accidents in one year				
Class interval	0	1	2	3	4 or more
18-25	600	260	185	90	70
25-40	900	240	160	85	65
40-50	1000	195	150	70	50
50 and above	500	170	120	60	30

If a driver from the city is selected at random, find the probability of the following events:

(i) Age lying between 18-25 and meet 2 accidents

(ii) Age between 25-50 and meet at least 3 accidents

(iii) Age more than 40 years and meet at most one accident

(iv) Having one accident in the year

(v) Having no accident in the year.

E 2) Past experience of 200 consecutive days speaks that weather forecasts of a station is 120 times correct. A day is selected at random of the year, find the probability that

(i) weather forecast on this day is correct

(ii) weather forecast on this day is false

E 3) Throw a die 200 times and find the probability of getting the odd number using statistical definition of probability.

2.4 SUBJECTIVE APPROACH TO PROBABILITY

In this approach, we try to assess the probability from our own experiences. This approach is applicable in the situations where the events do not occur at all or occur only once or cannot be performed repeatedly under the same conditions. Subjective probability is based on one's judgment, wisdom, intuition and expertise. It is interpreted as a measure of degree of belief or as the quantified judgment of a particular individual. For example, a teacher may express his /her confidence that the probability for a particular student getting first position in a test is 0.99 and that for a particular student getting failed in the test is 0.05. It is based on his personal belief.

You may notice here that since the assessment is purely subjective one, it will vary from person to person, depending on one's perception of the situation and past experience. Even when two persons have the same knowledge about the past, their assessment of probabilities may differ according to their personal prejudices and biases.

2.5 AXIOMATIC APPROACH TO PROBABILITY

All the approaches i.e. classical approach, relative frequency approach (Statistical/Empirical probability) and subjective approach share the same basic axioms. These axioms are fundamental to the probability and provide us with unified approach to probability i.e. axiomatic approach to probability. It defines the probability function as follows:

Let S be a sample space for a random experiment and A be an event which is subset of S, then P(A) is called probability function if it satisfies the following axioms

(i) P(A) is real and $P(A) \geq 0$

(ii) P (S) = 1

(iii) If $A_1, A_2, \ldots$ is any finite or infinite sequence of disjoint events in S, then

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

Now, let us give some results using probability function. But before taking up these results, we discuss some statements with their meanings in terms of set theory. If A and B are two events, then in terms of set theory, we write

i) ‘At least one of the events A or B occurs’ as $A \cup B$

ii) ‘Both the events A and B occurs’ as $A \cap B$

iii) ‘Neither A nor B occurs’ as $\bar{A} \cap \bar{B}$

iv) ‘Event A occurs and B does not occur’ as $A \cap \bar{B}$

v) ‘Exactly one of the events A or B occurs’ as $(\bar{A} \cap B) \cup (A \cap \bar{B})$

vi) ‘Not more than one of the events A or B occurs’ as

$$\left(A \cap \bar{B}\right) \cup \left(\bar{A} \cap B\right) \cup \left(\bar{A} \cap \bar{B}\right).$$

Similarly, you can write the meanings in terms of set theory for such statement in case of three or more events e.g. in case of three events A, B and C, happening of at least one of the events is written as $A \cup B \cup C$.

2.6 SOME RESULTS USING PROBABILITY FUNCTION

1 Prove that probability of the impossible event is zero

Proof: Let S be the sample space and ϕ be the set of impossible event.

$\therefore\ S \cup \phi = S$

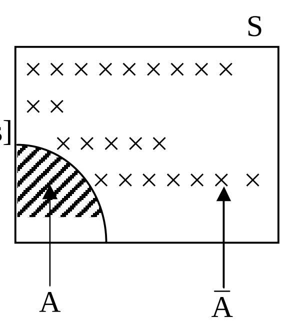
$\Rightarrow P(S \cup \phi) = P(S)$

$\Rightarrow P(S) + P(\phi) = P(S)$ [By axiom (iii)]

$\Rightarrow 1 + P(\phi) = 1$ [By axiom (ii)]

$\Rightarrow P(\phi) = 0$

2 Probability of non-happening of an event A i.e. complementary event $\bar{A}$ of A is given by $P(\bar{A}) = 1 - P(A)$



Proof: If S is the sample space then

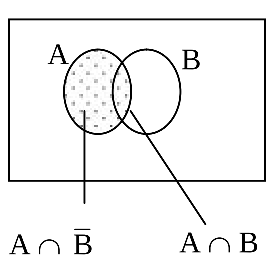
$A \cup \bar{A} = S$ [$\because$ A and $\bar{A}$ are mutually disjoint events]

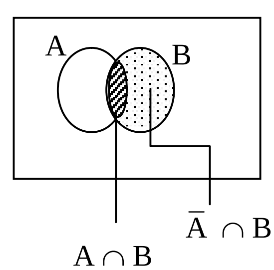
$\Rightarrow P(A \cup \bar{A}) = P(S)$

$\Rightarrow P(A) + P(\bar{A}) = P(S)$ [Using axiom (iii)]

$= 1$ [Using axiom (ii)]

$\Rightarrow P(\bar{A}) = 1 - P(A)$





3. Prove that

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

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

Proof

If S is the sample space and A, B $\subset$ S then

(i) $A = (A \cap \overline{B}) \cup (A \cap B)$

$\Rightarrow P(A) = P((A \cap \overline{B}) \cup (A \cap B))$

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

[Using axiom (iii) as $A \cap \overline{B}$ and $A \cap B$ are mutually disjoint]

$\Rightarrow P(A \cap \overline{B}) = P(A) - P(A \cap B)$

(ii) $B = (A \cap B) \cup (\overline{A} \cap B)$

$P(B) = P((A \cap B) \cup (\overline{A} \cap B))$

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

[Using axiom (iii) as $A \cap B$ and $\overline{A} \cap B$ are mutually disjoint]

$\Rightarrow P(\overline{A} \cap B) = P(B) - P(A \cap B)$

Example 4: A, B and C are three mutually exclusive and exhaustive events associated with a random experiment. Find P(A) given that :

$$P(B) = \frac{3}{4}P(A) \text{ and } P(C) = \frac{1}{3}P(B)$$

Solution:

As A, B and C are mutually exclusive and exhaustive events,

$\therefore A \cup B \cup C = S$

$\Rightarrow P(A \cup B \cup C) = P(S)$

$\Rightarrow P(A) + P(B) + P(C) = 1 \quad \left[\begin{array}{l}\because \text{ using axiom (iii) as A, B, C} \\ \text{are mutually disjoint events}\end{array}\right]$

$$\Rightarrow P(A) + \frac{3}{4}P(A) + \frac{1}{3}P(B) = 1$$

$$\Rightarrow P(A) + \frac{3}{4}P(A) + \frac{1}{3}\left(\frac{3}{4}P(A)\right) = 1$$

$$\Rightarrow \left(1 + \frac{3}{4} + \frac{1}{4}\right)P(A) = 1$$

$$\Rightarrow 2P(A) = 1$$

$$\Rightarrow P(A) = \frac{1}{2}$$

Examples 5: If two dice are thrown, what is the probability that sum is

a) greater than 9, and

b) neither 10 or 12.

Solution:

a) P[sum > 9] = P[sum = 10 or sum = 11 or sum = 12]

= P[sum =10] + P[sum = 11] + P[sum = 12]

[using axiom (iii)]

$$= \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}$$

[$\because$ for sum = 10, there are three favourable cases (4, 6), (5, 5) and (6, 4). Similarly for sum =11 and 12, there are two and one favourable cases respectively.]

Let A denotes the event for sum =10 and B denotes the event for sum = 12,

$\therefore$ Required probability = $P\left(\bar{A} \cap \bar{B}\right) = P\left(\overline{A \cup B}\right)$ [Using De- Morgan's law

(see Unit 1 of Course MST-001)]

$= 1 - P(A \cup B)$

$= 1 - [P(A) + P(B)]$ [Using axiom (iii)]

$$= 1 - \left[\frac{3}{36} + \frac{1}{36}\right] = 1 - \frac{4}{36} = 1 - \frac{1}{9} = \frac{8}{9}$$

Now, you can try the following exercises.

E4) If A, B and C are any three events, write down the expressions in terms of set theory:

a) only A occurs
b) A and B occur but C does not
c) A, B and C all the three occur
d) at least two occur
e) exactly two do not occur
f) none occurs

E5) Fourteen balls are serially numbered and placed in a bag. Find the probability that a ball is drawn bears a number multiple of 3 or 5.

2.7 SUMMARY

Let us summarize the main topics covered in this unit.

1) When classical definition fails, we obtain the probability by observing the data. This approach to probability is called the **relative frequency approach** and it defines the statistical probability. If an event A (say) happens m times in n trials of an experiment which is performed repeatedly under essentially homogeneous and identical conditions, then the (**Statistical or Empirical**) probability of happening A is defined as

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

2) **Subjective probability** is based on one's judgment, wisdom, intuition and expertise. It is interpreted as a measure of degree of belief or as the quantified judgment of particular individual.

3) If S be a sample space for a random experiment and A be an event which is subset of S, then P(A) is called probability function if it satisfies the following axioms

(i) P(A) is real and $P(A) \geq 0$

(ii) P (S) = 1

(iii) If $A_1, A_2, \ldots$ is any finite or infinite sequence of disjoint events in S, then

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

This is the **axiomatic approach to the probability**.

2.8 SOLUTIONS/ANSWERS

E 1) Since the information is based on 5000 drivers,

the number of exhaustive cases is = 5000.

Thus,

(i) the required probability = $\frac{185}{5000} = \frac{37}{1000}$

(ii) the required probability = $\frac{85+65+70+50}{5000} = \frac{270}{5000} = \frac{27}{500}$

(iii) the required probability = $\frac{1000+195+500+170}{5000} = \frac{1865}{5000} = \frac{373}{1000}$

(iv) the required probability = $\frac{260+240+195+170}{5000} = \frac{865}{5000} = \frac{173}{1000}$

(v) the required probability = $\frac{600+900+1000+500}{500} = \frac{3000}{5000} = \frac{3}{5}$

E 2) Since the information is based on the record of 200 days, so the number of exhaustive cases in each case = 200.

(i) Number of favourable cases for correct forecast = 120

$\therefore$ the required probability = $\frac{120}{200} = \frac{12}{20} = \frac{3}{5}$

(iii) Number of favourable outcomes for incorrect forecast = 200 – 120

$= 80$

$\therefore$ the required probability = $\dfrac{80}{200} = \dfrac{2}{5}$

E 3) First throw a die 200 times and note the outcomes. Then construct a table for the number of throws and the number of times the odd number turns up as shown in the following format:

Number of Throws(n)	Number of times the odd number turns up (m)	Proportion (m/n)
1		
2		
3		
.		
.		
.		
200		

Now, plot the graph taking number of throws (n) on x-axis and the proportion ($\frac{m}{n}$) on y-axis in the manner as shown in Fig. 2.1. Then see to which value the proportion ($\frac{m}{n}$) approaches to as n becoming large. This limiting value of $\frac{m}{n}$ is the required probability.

E 4) a) $A \cap \bar{B} \cap \bar{C}$

b) $A \cap B \cap \bar{C}$

c) $A \cap B \cap C$

d) $\left(A \cap B \cap \bar{C}\right) \cup \left(A \cap \bar{B} \cap C\right) \cup \left(\bar{A} \cap B \cap C\right) \cup \left(A \cap B \cap C\right)$

e) $\left(\bar{A} \cap \bar{B} \cap C\right) \cup \left(\bar{A} \cap B \cap \bar{C}\right) \cup \left(A \cap \bar{B} \cap \bar{C}\right)$

f) $\bar{A} \cap \bar{B} \cap \bar{C}$

E 5) Let A be the event that the drawn ball bears a number multiple of 3 and B be the event that it bears a number multiple of 5, then

A = {3, 6, 9, 12} and B = {5, 10}

$$\therefore \ P(A) = \frac{4}{14} = \frac{2}{7} \text{ and } P(B) = \frac{2}{14} = \frac{1}{7}$$

The required probability = P(A or B)

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

$$= \frac{2}{7} + \frac{1}{7} = \frac{3}{7}$$

[Using axiom (iii) as A and B are mutually disjoint]

UNIT 3 LAWS OF PROBABILITY

Structure

3.1 INTRODUCTION

In the preceding two units of this block, you have studied various approaches to probability, their direct applications and various types of events in terms of set theory. However, in many situations, we may need to find probability of occurrence of more complex events. Now, we are adequately equipped to develop the laws of probability i.e. the law of addition and the law of multiplication which will help to deal with the probability of occurrence of complex events.

Objectives

After completing this unit, you should be able to discuss:

- addition law of probability;
- conditional probability;
- multiplication law of probability;
- independent events;
- probability of happening at least one of the independent events; and
- problems on addition and multiplicative laws of probability.

3.2 ADDITION LAW

Addition Theorem on Probability for Two Events

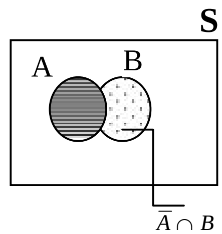
Statement

Let S be the sample space of a random experiment and events A and $B \subseteq S$ then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof: From the Venn diagram, we have

$$A \cup B = A \cup \left(\bar{A} \cap B\right)$$

$$\therefore\ P\left(A \cup B\right) = P\left(A\right) + P\left(\bar{A} \cap B\right) \qquad \left[\begin{array}{l}\text{By axiom (iii)} \because\ A \text{ and } \bar{A} \cap B \\ \text{are mutually disjoint}\end{array}\right]$$



$$= P(A) + P(B) - P\left(A \cap B\right) \left[\begin{array}{l}\text{Refer to result 3 of Sec. 2.6} \\ \text{of Unit 2 of this Course}\end{array}\right]$$

Hence proved

Corollary: If events A and B are mutually exclusive events, then

$P(A \cup B) = P(A) + P(B)$.

[This is known as the addition theorem for mutually exclusive events]

Proof: For any two events A and B, we know that

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

Now, if the events A and B are mutually exclusive then

$A \cap B = \phi$

Also, we know that probability of impossible event is zero i.e.

$P(A \cap B) = P(\phi) = 0$.

Hence,

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

Similarly, for three non-mutually exclusive events A, B and C, we have

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

and for three mutually exclusive events A, B and C, we have

$$P\left(A \cup B \cup C\right) = P\left(A\right) + P\left(B\right) + P\left(C\right).$$

The result can similarly be extended for more than 3 events.

Applications of Addition Theorem of Probability

Example 1: From a pack of 52 playing cards, one card is drawn at random. What is the probability that it is a jack of spade or queen of heart?

Solution: Let A and B be the events of drawing a jack of spade and queen of heart, respectively.

$$\therefore\ P(A) = \frac{1}{52} \text{ and } P(B) = \frac{1}{52} \begin{bmatrix} \because \text{there is one card each of jack} \\ \text{of spade and queen of heart} \end{bmatrix}$$

Here, a card cannot be both the jack of spade and the queen of heart, hence A and B are mutually exclusive,

$\therefore$ applying the addition theorem for mutually exclusive events,

the required probability = $P(A \cup B) = P(A) + P(B)$

$$= \frac{1}{52} + \frac{1}{52} = \frac{2}{52} = \frac{1}{26}.$$

Example 2: 25 lottery tickets are marked with first 25 numerals. A ticket is drawn at random. Find the probability that it is a multiple of 5 or 7.

Solution: Let A be the event that the drawn ticket bears a number multiple of 5 and B be the event that it bears a number multiple of 7.

Therefore,

A = {5, 10, 15, 20, 25}

B = {7, 14, 21}

Here, as $A \cap B = \phi$,

$\therefore$ A and B are mutually exclusive, and hence,

$$P(A \cup B) = P(A) + P(B) = \frac{5}{25} + \frac{3}{25} = \frac{8}{25}$$

Example 3: Find the probability of getting either a number multiple of 3 or a prime number when a fair die is thrown.

Solution: When a die is thrown, then the sample space is

S = {1, 2, 3, 4, 5, 6}

Let A be the event of getting a number multiple of 3 and B be the event of getting a prime number,

$\therefore A = \{3, 6\}$, $B = \{2, 3, 5\}$, $A \cap B = \{3\}$

Here as $A \cap B$ is not empty set,

$\therefore$ A and B are non-mutually exclusive and hence,

the required probability = $P(A \cup B)$

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

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

$$= \frac{2+3-1}{6} = \frac{4}{6} = \frac{2}{3}.$$

Example 4: There are 40 pages in a book. A page is opened at random. Find the probability that the number of this opened page is a multiple of 3 or 5.

Solution Let A be the event that the number of the opened page is a multiple of 3 and B be the event that it is a multiple of 5.

$\therefore$ A = {3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39},

B = {5, 10, 15, 20, 25, 30, 35, 40}, and

$A \cap B = \{15, 30\}$

As A and B are non-mutually exclusive,

$\therefore$ the required probability = $P(A \cup B)$

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

$$= \frac{13}{40} + \frac{8}{40} - \frac{2}{40} = \frac{19}{40}.$$

Example 5: A Card is drawn from a pack of 52 playing cards, find the proabability that the drawn card is an ace or a red colour card.

Solution: Let A be the event that the drawn card is a card of ace and B be the event that it is red colour card.

Now as there are four cards of ace and 26 red colour cards in a pack of 52 playing cards. Also, 2 cards in the pack are ace cards of red colour.

$$\therefore P(A) = \frac{4}{52},\ P(B) = \frac{26}{52}\ \text{, and}\ P(A \cap B) = \frac{2}{52}$$

$\therefore$ the required probability = $P(A \cup B)$

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

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52} = \frac{7}{13}.$$

Now, you can try the following exercises.

E1) A card is drawn from a pack of 52 playing cards. Find the probability that it is either a king or a red card.

E2) Two dice are thrown together. Find the probability that the sum of the numbers turned up is either 6 or 8.

3.3 CONDITIONAL PROBABILITY AND MULTIPLICATIVE LAW

Conditional Probability

We have discussed earlier that P(A) represents the probability of happening event A for which the number of exhaustive cases is the number of elements in the sample space S. P(A) dealt earlier was the unconditional probability. Here, we are going to deal with conditional probability.

Let us start with taking the following example:

Suppose a card is drawn at random from a pack of 52 playing cards. Let A be the event of drawing a black colour face card. Then $A = \{J_s, Q_s, K_s, J_c, Q_c, K_c\}$ and hence

P(A) = 6/52 = 3/26.

Let B be the event of drawing a card of spade i.e.

$B = \{1_s, 2_s, 3_s, 4_s, 5_s, 6_s, 7_s, 8_s, 9_s, 10_s, J_s, Q_s, K_s\}$.

If after a card is drawn from the pack of cards, we are given the information that card of spade has been drawn i.e., B has happened, then the probability of event A will no more be $\frac{3}{26}$, because here in this case, we have the information that the card drawn is of spade (i.e. from amongst 13 cards) and hence there are 13 exhaustive cases and not 52. From amongst these 13 cards of spade, there are 3 black colour face cards and hence probability of having black colour face card given that it is a card of spade i.e. $P(A|B) = 3/13$, which is the conditional probability of A given that B has already happened.

Note: Here, the symbol '|' used in P(A|B) should be read as 'given' and not 'upon'. P(A|B) is the conditional probability of happening A given that B has already happened i.e. here A happens depending on the condition of B.

So, the conditional probability P(A|B) is also the probability of happening A but here the information is given that the event B has already happened. P(A|B) refers to the sample space B and not S.

Remark 1: P(A|B) is meaningful only when $P(B) \neq 0$ i.e. when the event B is not an impossible event.

Multiplication Law of Probability

Statement: For two events A and B,

$$P(A \cap B) = P(A)\, P(B|A), \quad P(A) > 0 \qquad \ldots (1)$$

$$= P(B)\, P(A|B), \quad P(B) > 0, \qquad \ldots (2)$$

where P(B|A) is the conditional probability of B given that A has already happened and P(A|B) is the conditional probability of A given that B has already happened.

Proof: Let n be the number of exhaustive cases corresponding to the sample space S and m_1, m_2, m_3 be the number of favourable cases for events A, B and $A \cap B$ respectively.

$$\therefore \; P(A) = \frac{m_1}{n}, \; P(B) = \frac{m_2}{n}, \; P(A \cap B) = \frac{m_3}{n}$$

Now, as B|A represents the event of happening B given that A has already happened and hence it refers to the sample space A ($\because$ we have with us the information that A has already happened) and thus the number of exhaustive cases for B|A is m_1 (i.e. the number of cases favourable to "A relative to sample space S"). The number of cases favourable to B|A is the number of those elements of B which are in A i.e. the number of favourable cases to B|A is the number of favourable cases to "$A \cap B$ relative to S". So, the number of favourable cases to B|A is m_3

$$\therefore \; P(B|A) = \frac{m_3}{m_1}$$

$$= \frac{m_3/n}{m_1/n} \qquad \begin{bmatrix} \text{Dividing the numerator} \\ \text{and denominator by n} \end{bmatrix}$$

$$= \frac{P(A \cap B)}{P(A)},\ P(A) \neq 0$$

$\Rightarrow P(A \cap B) = P(A)\ P(B|A),\ P(A) \neq 0$

Similarly, you can prove yourself that

$P(A \cap B) = P(B)\ P(A|B),\ P(B) \neq 0.$

This result can be extended for 3 or more events e.g. for three events A, B and C, we have

$P(A \cap B \cap C) = P(A)\ P(B|A)\ P(C|A \cap B),$

where $P(C|A \cap B)$ represents the probability of happening C given that A and B both have already happened.

3.4 INDEPENDENT EVENTS

Before defining the independent events, let us again consider the concept of conditional probability taking the following example:

Suppose, we draw a card from a pack of 52 playing cards, then probability of drawing a card of spade is 13/52. Now, if we do not replace the card back and draw the next card. Then, the probability of drawing the second card 'a card of spade' if it being given that the first card was spade would be 12/51 and it is the conditional probability. Now, if the first card had been replaced back then this conditional probability would have been 13/52. So, if sampling is done without replacement, the probability of second draw and that of subsequent draws made following the same way is affected but if it is done with replacement, then the probability of second draw and subsequent draws made following the same way remains unaltered.

So, if in the above example, if the next draw is made with replacement, then the happening or non-happening of any draw is not affected by the preceding draws. Let us now define independent events.

Independent Events

Events are said to be independent if happening or non-happening of any one event is not affected by the happening or non-happening of other events. For example, if a coin is tossed certain number of times, then happening of head in any trial is not affected by any other trial i.e. all the trials are independent.

Two events A and B are independent if and only if $P(B|A) = P(B)$ i.e. there is no relevance of giving any information. Here, if A has already happened, even then it does not alter the probability of B. e.g. Let A be the event of getting head in the 4^{th} toss of a coin and B be the event of getting head in the 5^{th} toss of the coin. Then the probability of getting head in the 5^{th} toss is $\frac{1}{2}$, irrespective of the case whether we know or don't know the outcome of 4^{th} toss, i.e. $P(B|A) = P(B)$.

Multiplicative Law for Independent Events:

If A and B are independent events, then

$P(A \cap B) = P(A)\ P(B).$

This is because if A and B are independent then $P(B|A) = P(A)$ and hence the equation (1) discussed in Sec. 3.3 of this unit becomes $P(A \cap B) = P(A)\, P(B)$.

Similarly, if A, B and C are three independent events, then

$P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$.

The result can be extended for more than three events also.

Remark 2: Mutually exclusive events can never be independent.

Proof: Let A and B be two mutually exclusive events with positive probabilities (i.e. $P(A) > 0$, and $P(B) > 0$)

$\therefore\ P(A \cap B) = 0\ [\because A \cap B = \phi]$

Also, by multiplication law of probability, we have

$P(A \cap B) = P(A)\, P(B|A),\ P(A) \neq 0.$

$\therefore\ 0 = P(A)\, P(B|A).$

Now as $P(A) \neq 0$, $\therefore\ P(B|A) = 0$

But $P(B) \neq 0 \quad [\because P(B) > 0]$

$\therefore\ P(B|A) \neq P(B)$,

Hence A and B are not independent.

Result: If events A and B are independent then prove that

(i) A and $\overline{B}$ are independent

(ii) $\overline{A}$ and B are independent

(iii) $\overline{A}$ and $\overline{B}$ are independent

Proof:

(i) We know that

$$\begin{aligned} P(A \cap \overline{B}) &= P(A) - P(A \cap B) && \text{[Already proved in Unit 2 of this course]} \\ &= P(A) - P(A)P(B) && [\because \text{events A and B are independent.}] \\ &= P(A)(1 - P(B)) \\ &= P(A)P(\overline{B}) \end{aligned}$$

$\Rightarrow$ Events A and $\overline{B}$ are independent

(ii) We know that

$$\begin{aligned} P(\overline{A} \cap B) &= P(B) - P(A \cap B) && \text{[Already proved in Unit 2 of this course]} \\ &= P(B) - P(A)P(B) && [\because \text{ events A and B are independent.}] \\ &= P(B)(1 - P(A)) \\ &= P(B)P(\overline{A}) \\ &= P(\overline{A})P(B) \end{aligned}$$

$\Rightarrow$ Events $\overline{A}$ and B are independent

(iii) $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B})$ [By De-Morgan's law]

$$= 1 - P(A \cup B) \quad \left[\because P(E) + P(\overline{E}) = 1\right]$$

$$= 1 - [P(A) + P(B) - P(A \cap B)] \text{ [Using addition law on probability]}$$

$$= 1 - [P(A) + P(B) - P(A)P(B)] \left[\because \text{events A and B are independent.}\right]$$

$$= 1 - P(A) - P(B) + P(A)P(B)$$

$$= (1 - P(A)) - P(B)(1 - P(A))$$

$$= (1 - P(A))(1 - P(B))$$

$$= P(\overline{A})P(\overline{B})$$

$\Rightarrow$ Events $\overline{A}$ and $\overline{B}$ are independent

Now let us take up some examples on conditional probability, multiplicative law and independent events:

Example 6: A die is rolled. If the outcome is a number greater than 3, what is the probability that it is a prime number?

Solution: The sample space of the experiment is

S = {1, 2, 3, 4, 5, 6}

Let A be event that the outcome is a number greater than 3 and B be the event that it is a prime number.

$\therefore$ A = {4, 5, 6}, B = {2, 3, 5} and hence $A \cap B$ = {5}.

$\Rightarrow$ P(A) = 3/6, P(B) = 3/6, $P(A \cap B)$ = 1/6.

Now, the required probability = P(B|A)

$$= \frac{P(A \cap B)}{P(A)} \quad \begin{bmatrix} \text{Refer the multiplication} \\ \text{law given by (1) in} \\ \text{Sec. 3.3 of the Unit} \end{bmatrix}$$

$$= \frac{1/6}{3/6} = \frac{1}{3}$$

Example 7: A couple has 2 children. What is the probability that both the children are boys, if it is known that?

(i) younger child is a boy

(ii) older child is a boy

(iii) at least one of them is boy

Solution: Let B_i, G_i denote that i^{th} birth is of boy and girl respectively, i =1, 2.

Then for a couple having two children, the sample space is

$S = \{B_1B_2, B_1G_2, G_1B_2, G_1G_2\}$

Let A be the event that both children are boys then

$A = \{B_1B_2\}$

(i) Let B be the event of getting younger child as boy i.e.

$B = \{B_1B_2, G_1B_2\}$. Hence $A \cap B = \{B_1B_2\}$

$$\therefore \text{required probability } P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{2/4} = \frac{1}{2}$$

(ii) Let C be the event of getting older child as boy, then $C = \{B_1B_2, B_1G_2\}$

and hence $A \cap C = \{B_1B_2\}$.

$$\therefore \text{ required probability } P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{1/4}{2/4} = \frac{1}{2}.$$

(iii) Let D be the event of getting at least one of the children as boy, then

$D = \{B_1B_2, B_1G_2, G_1B_2\}$ and hence

$A \cap D = \{B_1B_2\}$.

$$\therefore \text{required probability } P(A \mid D) = \frac{P(A \cap D)}{P(D)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Example 8: An urn contains 4 red and 7 blue balls. Two balls are drawn one by one without replacement. Find the probability of getting 2 red balls.

Solution: Let A be the event that first ball drawn is red and B be the event that the second ball drawn is red.

$$\therefore P(A) = 4/11 \text{ and } P(B|A) = 3/10 \left[\begin{array}{l}\because \text{it is given that one red ball} \\ \text{has already been drawn}\end{array}\right]$$

$$\therefore \text{The required probability} = P(A \text{ and } B)$$

$$= P(A)\,P(B|A)$$

$$= \left(\frac{4}{11}\right)\left(\frac{3}{10}\right) = \frac{6}{55}$$

Example 9: Three cards are drawn one by one without replacement from a well shuffled pack of 52 playing cards. What is the probability that first card is jack, second is queen and the third is again a jack.

Solution: Define the following events

E_1 be the event of getting a jack in the first draw,

E_2 be the event of getting a queen in second draw, and

E_3 be the event of getting a jack in third draw,

$\therefore$ Required probability = $P(E_1 \cap E_2 \cap E_3)$

$$= P(E_1)P(E_2|E_1)P(E_3 | E_1 \cap E_2)$$

$$\left[\begin{array}{l}\because \text{ cards are drawn without replacement and} \\ \text{hence the events are not independent}\end{array}\right]$$

$$= \frac{4}{52} \times \frac{4}{51} \times \frac{3}{50} = \frac{1}{13} \times \frac{2}{17} \times \frac{1}{25} = \frac{2}{5525}.$$

Example 10: (i) If A and B are independent events with

$P(A \cup B) = 0.8$ and $P(B) = 0.4$ then find $P(A)$.

(ii) If A and B are independent events with

$P(A) = 0.2$, $P(B) = 0.5$ then find $P(A \cup B)$.

(iii) If A and B are independent events and

$P(A) = 0.4$ and $P(B) = 0.3$, then find $P(A|B)$ and $P(B|A)$.

(iv) If A and B are independent events with $P(A) = 0.4$ and $P(B) = 0.2$, then find

$P(\overline{A} \cap B)$, $P(A \cap \overline{B})$, $P(\overline{A} \cap \overline{B})$

Solution:

(i) We are given

$P(A \cup B) = 0.8$, $P(B) = 0.4$.

We know that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{[By Addition theorem of probability]}$$

$$= P(A) + P(B) - P(A)P(B) \quad [\because \text{ events A and B are independent }]$$

$$\Rightarrow \; 0.8 = P(A) + 0.4 - 0.4P(A)$$

$$\Rightarrow \; 0.4 = (1 - 0.4)\, P(B)$$

$$= 0.6\, P(B)$$

$$\Rightarrow P(B) = \frac{0.4}{0.6} = \frac{4}{6} = \frac{2}{3}.$$

(ii) We are given that $P(A) = 0.2$, $P(B) = 0.5$.

We know that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{[Addition theorem of probability]}$$

$$= P(A) + P(B) - P(A)P(B) \quad [\because \text{ events A and B are independent}]$$

$$= 0.2 + 0.5 - 0.2 \times 0.5$$

$$= 0.7 - 0.10 = 0.6$$

(iii) We are given that $P(A) = 0.4$, $P(B) = 0.3$.

Now, as A and B are independent events,

$\therefore P(A \cap B) = P(A)P(B)$

$= 0.4 \times 0.3 = 0.12$

And hence from conditional probability, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.12}{0.3} = \frac{12}{30} = \frac{2}{5} = 0.4,$$

and $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.12}{0.4} = \dfrac{12}{40} = \dfrac{3}{10} = 0.3$.

(iv) We are given that P(A) = 0.4, P(B) = 0.2.

We know that if two events A and B are independent then

$\bar{A}$ and B; A and $\bar{B}$; $\bar{A}$ and $\bar{B}$ are also independent events.

$\therefore P(\bar{A} \cap B) = P(\bar{A})P(B)$ [Using the concept of independent events]

$= (1 - P(A))(P(B))$

$= (1 - 0.4)(0.2)$

$= (0.6)(0.2)$

$= 0.12$

$P(A \cap \bar{B}) = P(A)\, P(\bar{B})$ [$\because$ A and $\bar{B}$ are independent]

$= (0.4)(1 - 0.2) = 0.32$

$$P\left(\bar{A} \cap \bar{B}\right) = P\left(\bar{A}\right)P\left(\bar{B}\right) = \left(1-0.4\right)\left(1-0.2\right) = \left(0.6\right)\left(0.8\right) = 0.48.$$

Example 11: Three unbiased coins are tossed simultaneously. In which of the following cases are the events A and B independent?

(i) A be the event of getting exactly one head

B be the event of getting exactly one tail

(ii) A be the event that first coin shows head

B be the event that third coin shows tail

(iii) A be the event that shows exactly two tails

B be the event that third coin shows head

Solution: When three unbiased coins are tossed simultaneously, then the sample space is given by

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

(i) A = {HTT, THT, TTH}

B = {HHT, HTH, THH}

$A \cap B = \{\} = \phi$

$\therefore P(A) = \dfrac{3}{8}, P(B) = \dfrac{3}{8}, P(A \cap B) = \dfrac{0}{8} = 0$

Hence, P(A) P(B) = $\dfrac{3}{8} \times \dfrac{3}{8} = \dfrac{9}{64}$

$\Rightarrow P(A \cap B) \neq P(A)P(B)$

$\Rightarrow$ Events A and B are not independent.

(ii) A = {HHH, HHT, HTH, HTT}

B = {HHT, HTT, THT, TTT}

$A \cap B$ = {HHT, HTT}

$\therefore P(A) = \frac{4}{8} = \frac{1}{2}$, $P(B) = \frac{4}{8} = \frac{1}{2}$, $P(A \cap B) = \frac{2}{8} = \frac{1}{4}$

$P(A)P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = P(A \cap B)$

$\Rightarrow$ Events A and B are independent.

(iii) A= {HTT, THT, TTH}

B = {HHH, HTH, THH, TTH}

$A \cap B$ = {TTH}

$P(A) = \frac{3}{8}$, $P(B) = \frac{4}{8} = \frac{1}{2}$, $P(A \cap B) = \frac{1}{8}$.

Hence, $P(A)P(B) = \frac{3}{8} \times \frac{1}{2} = \frac{3}{16} \neq P(A \cap B)$

$\Rightarrow$ Events A and B are not independent.

Example 12: Two cards are drawn from a pack of cards in succession with replacement of first card. Find the probability that both are the cards of 'heart'.

Solution: Let A be the event that the first card drawn is a heart card and B be the event that second card is a heart card.

As the cards are drawn with replacement,

$\therefore$ A and B are independent and hence the required probability

$$= P(A \cap B) = P(A)P(B) = \left(\frac{13}{52}\right)\left(\frac{13}{52}\right) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}.$$

Example 13: A class consists of 10 boys and 40 girls. 5 of the students are rich and 15 students are brilliant. Find the probability of selecting a brilliant rich boy.

Solution: Let A be the event that the selected student is brilliant, B be the event that he/she is rich and C be the event that the student is boy.

$\therefore P(A) = \frac{15}{50}, P(B) = \frac{5}{50}, P(C) = \frac{10}{50}$ and hence

the required probability = $P(A \cap B \cap C)$

$$= P(A)P(B)P(C) \quad [\because \text{A, B and C are independent}]$$

$$= \left(\frac{15}{50}\right)\left(\frac{5}{50}\right)\left(\frac{10}{50}\right) = \left(\frac{3}{10}\right)\left(\frac{1}{10}\right)\left(\frac{1}{5}\right) = \frac{3}{500}$$

Here are some exercises for you.

E3) A card is drawn from a well-shuffled pack of cards. If the card drawn is a face card, what is the probability that it is a king?

E4) Two cards are drawn one by one without replacement from a well shuffled pack of 52 cards. What is the probability that both the cards are red?

E5) A bag contains 10 good and 4 defective items, two items are drawn one by one without replacement. What is the probability that first drawn item is defective and the second one is good?

E6) The odds in favour of passing driving test by a person X are $3:5$ and odds in favour of passing the same test by another person Y are $3:2$. What is the probability that both will pass the test?

3.5 PROBABILITY OF HAPPENING AT LEAST ONE OF THE INDEPENDENT EVENTS

If A and B be two independent events, then probability of happening at least one of the events is

$$P(A \cup B) = 1 - P(\overline{A \cup B})$$

$$= 1 - P(\overline{A} \cap \overline{B}) \qquad \text{[By DeMorgan's Law]}$$

$$= 1 - P(\overline{A})P(\overline{B}) \qquad \left[\because \text{ A and B and hence } \overline{A} \text{ and } \overline{B} \text{ are independent.}\right]$$

Similarly if we have n independent event $A_1, A_2, \ldots, A_n$, then probability of happening at least one of the events is

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = 1 - \left[P(\overline{A}_1)P(\overline{A}_2) \ldots P(\overline{A}_n)\right]$$

I.e. probability of happening at least one of the independent events

= 1 – probability of happening none of the events.

Example 14: A person is known to hit the target in 4 out of 5 shots whereas another person is known to hit 2 out of 3 shots. Find the probability that the target being hit when they both try.

Solution: Let A be the event that first person hits the target and B be the event that second person hits the target.

$$\therefore P(A) = \frac{4}{5}, \quad P(B) = \frac{2}{3}$$

Now, as both the persons try independently,

$\therefore$ the required probability = probability that the target is hit

= probability that at least one of the persons hits the target

$= P(A \cup B)$

$= 1 - P(\overline{A})P(\overline{B})$

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

$$= 1 - \left(\frac{1}{5}\right)\left(\frac{1}{3}\right) = 1 - \frac{1}{15} = \frac{14}{15}.$$

E 7) A problem in statistics is given to three students A, B and C whose chances of solving it are 0.3, 0.5 and 0.6 respectively. What is the probability that the problem is solved?

3.6 PROBLEMS USING BOTH ADDITION AND MULTIPLICATIVE LAWS

Here we give some examples which are based on both the addition and multiplication laws.

Example15: Husband and wife appear in an interview for two vacancies for the same post. The probabilities of husband's and wife's selections are $\frac{2}{5}$ and $\frac{1}{5}$ respectively. Find the probability that

(i) Exactly one of them is selected

(ii) At least one of them is selected

(iii) None is selected.

Solution: Let H be the event that husband is selected and W be the event that wife is selected. Then,

$$P(H) = \frac{2}{5}, \; P(W) = \frac{1}{5}$$

$$\therefore P(\overline{H}) = 1 - \frac{2}{5} = \frac{3}{5}, \; P(\overline{W}) = 1 - \frac{1}{5} = \frac{4}{5}$$

(i) The required probability $= P\left[(H \cap \overline{W}) \cup (\overline{H} \cap W)\right]$

$$= P(H \cap \overline{W}) + P(\overline{H} \cap W)$$

$\left[\begin{array}{l}\text{By Addition theorem for} \\ \text{mutually exclusive events}\end{array}\right]$

$$= P(H)P(\overline{W}) + P(\overline{H})P(W)$$

$\left[\begin{array}{l}\because \text{selection of husband and} \\ \text{wife are independent}\end{array}\right]$

$$= \frac{2}{5} \times \frac{4}{5} + \frac{3}{5} \times \frac{1}{5} = \frac{8}{25} + \frac{3}{25} = \frac{11}{25}.$$

(ii) The required probability $= P(H \cup W)$

$= 1 - P\left(\overline{H}\right)P\left(\overline{W}\right)$

$$\left[\begin{array}{l}\because \text{H and W are independent and hence} \\ \overline{\text{H}}\text{ and }\overline{\text{W}}\text{ are independent events}\end{array}\right]$$

$$= 1-\frac{3}{5}\times\frac{4}{5} = 1-\frac{12}{25}=\frac{13}{25}.$$

(iii) The required probability = $P(\overline{H}\cap\overline{W})$

$$= P(\overline{H})P(\overline{W}) \left[\because \overline{H} \text{ and } \overline{W} \text{ are independent}\right]$$

$$=\frac{3}{5}\times\frac{4}{5}=\frac{12}{25}$$

Example 16: A person X speaks the truth in 80% cases and another person Y speaks the truth in 90% cases. Find the probability that they contradict each other in stating the same fact.

Solution: Let A, B be the events that person X and person Y speak truth respectively, then

$$P(A)=\frac{80}{100}=0.8,\ P(B) = \frac{90}{100}=0.9.$$

$\therefore P(\overline{A}) = 1-0.8 = 0.2,\ P(\overline{B}) = 1-0.9 = 0.1.$

Thus, the required probability $= P\left[\left(A\cap\overline{B}\right)\cup\left(\overline{A}\cap B\right)\right]$

$$= P(A\cap\overline{B})+P(\overline{A}\cap B) \left[\begin{array}{l}\text{By addition law for mutually} \\ \text{exclusive events}\end{array}\right]$$

$$= P(A)P(\overline{B})+P(\overline{A})P(B) \left[\begin{array}{l}\text{By multiplication law for} \\ \text{independent events}\end{array}\right]$$

$= 0.8 \times 0.1 + 0.2 \times 0.9$

$= 0.08+ 0.1= 0.26 = 26\%.$

Here is an exercise for you.

E 8) Two cards are drawn from a pack of cards in succession presuming that drawn cards are replaced. What is the probability that both drawn cards are of the same suit?

3.7 SUMMARY

Let us summarize the main points covered in this unit:

1) For two non-mutually exclusive events A and B, the **addition law** of probability is given by $P(A\cup B)=P(A)+P(B)-P(A\cap B)$. If these events are mutually exclusive then, $P(A\cup B)=P(A)+P(B)$.

2) **Conditional probability** for happening of an event say B given that an event A has already happened is given by

$P(B|A) = \dfrac{P(A \cap B)}{P(A)}$, P(A) > 0. The **Multiplicative law** of probability for any two events is stated as $P(A \cap B) = P(A)\,P(B|A)$.

3) Events are said to be **independent** if happening or non-happening of any one event is not affected by the happening or non-happening of other events. Two events A and B are independent if and only if $P(B|A) = P(B)$ and hence **multiplicative law for two independent events** is given by $P(A \cap B) = P(A)\,P(B)$.

4) If events are independent then their complements are also independent. **Probability of happening at least one of the independent events** can be obtained on subtracting from 1 the probability of happening none of the events.

3.8 SOLUTIONS/ANSWERS

E1) Let A be the event of getting a card of king and B be the event of getting a red card.

$\therefore$ the required probability = $P(A \cup B)$

$$= P(A) + P(B) - P(A \cap B) \quad \text{[By addition theorem]}$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} \quad \left[\begin{array}{l}\because \text{there are two cards which} \\ \text{are both king as well as red]}\end{array}\right]$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52}$$

$$= \frac{4+26-2}{52} = \frac{28}{52} = \frac{7}{13}$$

E2) Here, the number of exhaustive cases = $6 \times 6 = 36$.

Let A be the event that the sum is 6 and B be the event that the sum is 8.

$\therefore$ A = {(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)}, and

B = {(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)}.

Here, as $A \cap B = \phi$,

$\therefore$ A and B are mutually exclusive and hence

the required probability = $P(A \cup B) = P(A) + P(B)$

$$= \frac{5}{36} + \frac{5}{36} = \frac{10}{36} = \frac{5}{18}$$

E 3) Let A be the event that the card drawn is face card and B be the event that it is a king.

$$\therefore\ A = \{J_s, Q_s, K_s, J_h, Q_h, K_h, J_d, Q_d, K_d, J_c, Q_c, K_c\}$$

$B = \{K_s, K_h, K_c, K_d\}$

$\Rightarrow A \cap B = \{K_s, K_h, K_c, K_d\}$

$\Rightarrow P(A) = \frac{12}{52},\ P(B) = \frac{4}{52},$ and $P(A \cap B) = \frac{4}{52}$.

$\therefore$ the required probability $= P(B\,|\,A) = \frac{P(A \cap B)}{P(A)} = \frac{4/52}{12/52} = \frac{4}{12} = \frac{1}{3}$.

E 4) Let A be the event that first card drawn is red and B be the event that second card is red.

$\therefore P(A) = \frac{26}{52}$ and $P(B|A) = \frac{25}{51}$.

Thus, the desired probability $= P(A \cap B)$

$$= P(A)\,P(B|A) \quad \begin{bmatrix}\text{By multiplication}\\ \text{theorem}\end{bmatrix}$$

$$= \left(\frac{26}{52}\right)\left(\frac{25}{51}\right) = \frac{25}{102}.$$

E 5) Let E_1 be the event of getting a defective item in first draw and E_2 be the event of getting a good item in the second draw.

$\therefore$ the required probability $= P(E_1 \cap E_2)$

$$= P(E_1)P(E_2\,|\,E_1) \quad \begin{bmatrix}\text{By multiplication}\\ \text{theorem}\end{bmatrix}$$

$$= \frac{4}{14} \times \frac{10}{13} = \frac{20}{91}.$$

E 6) Let A be the event that person X passes the test and B be the event that person Y passes the test.

$\therefore P(A) = \frac{3}{3+5} = \frac{3}{8}$

and $P(B) = \frac{3}{3+2} = \frac{3}{5}$

Now, as both the person take the test independently,

$\therefore$ the required probability $= P(A \cap B)$

$$= P(A)P(B) = \left(\frac{3}{8}\right)\left(\frac{3}{5}\right) = \frac{9}{40}$$

E 7) Let E_1, E_2 and E_3 be the events that the students A, B and C solves the problem respectively.

$\therefore P(E_1) = 0.3,\ P(E_2) = 0.5$ and $P(E_3) = 0.6$.

Now, as the students try independently to solve the problem,

$\therefore$ the probability that the problem will be solved

= Probability that at least one of the students solves the problem

$= P(E_1 \cup E_2 \cup E_3)$

$= 1 - P\left(\bar{E}_1\right)P\left(\bar{E}_2\right)P\left(\bar{E}_3\right)$

$= 1 - (1 - 0.3)\,(1 - 0.5)\,(1 - 0.6)$

$= 1 - (0.7)\,(0.5)\,(0.4)$

$= 1 - 0.140 = 0.86.$

E 8) Let S_1, C_1, H_1 and D_1 be the events that first card is of spade, club, heart and diamond respectively; and let S_2, C_2, H_2 and D_2 be the events that second card is of spade, club, heart and diamond respectively.

Thus, the required probability

$= P[(S_1 \cap S_2) \text{ or } (C_1 \cap C_2) \text{ or } (H_1 \cap H_2) \text{ or } (D_1 \cap D_2)]$

$= P(S_1 \cap S_2) + P(C_1 \cap C_2) + P(H_1 \cap H_2) + P(D_1 \cap D_2)$

$= P(S_1)P(S_2) + P(C_1)P(C_2) + P(H_1)P(H_2) + P(D_1)P(D_2)$

$\left[\begin{array}{l}\because \text{cards are drawn with replacement and} \\ \text{hence the draws are independent}\end{array}\right]$

$= \frac{13}{52} \times \frac{13}{52} + \frac{13}{52} \times \frac{13}{52} + \frac{13}{52} \times \frac{13}{52} + \frac{13}{52} \times \frac{13}{52}$

$= 4 \times \frac{13}{52} \times \frac{13}{52}$

$= \frac{1}{4}$

UNIT 4 BAYES' THEOREM

Structure

4.1 INTRODUCTION

In the first three units of this block, we have seen as to how probability of different types of events is calculated. We have also discussed the operation of events and evaluation of their probabilities by using addition and multiplication laws. Still, there are situations whose probability evaluations require use of more results.

Probabilities obtained in the earlier units are per-revised or priori probabilities. However, these probabilities can be revised on the basis of some new related information. The revised probabilities are obtained using Bayes' theorem for which knowledge of total probability is also required. So, the present unit first discusses the law of total probability, its applications and then Bayes' theorem and its applications.

Objectives

After completing this unit, you should be able to:

- explain law of total probability;
- know as to how and when to apply the law of total probability;
- learn Bayes' theorem; and
- learn as to how and when to apply Bayes' theorem.

4.2 LAW OF TOTAL PROBABILITY

There are experiments which are conducted in two stages for completion. Such experiments are termed as two-stage experiments. At the first stage, the experiment involves selection of one of the given numbers of possible mutually exclusive events. At the second stage, the experiment involves happening of an event which is a sub-set of at least one of the events of first stage.

As an illustration for a two-stage experiment, let us consider the following example:

Suppose there are two urns – Urn I and Urn II. Suppose Urn I contains 4 white, 6 blue and Urn II contains 4 white, 5 blue balls. One of the urns is selected at random and a ball is drawn. Here, the first stage is the selection of one of the urns and second stage is the drawing of a ball of particular colour.

If we are interested in finding the probability of the event of second stage, then it is obtained using law of total probability, which is stated and proved as under:

Law of Total Probability

Statement: Let S be the sample space and $E_1, E_2, \ldots, E_n$ be n mutually exclusive and exhaustive events with $P(E_i) \neq 0$; $i = 1, 2, \ldots, n$. Let A be any event which is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$ (i.e. at least one of the events $E_1, E_2, \ldots, E_n$) with $P(A) > 0$, then $P(A) = P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2) + \ldots + P(E_n)\, P(A|E_n)$

$$= \sum_{i=1}^{n} P\left(E_i\right) P\left(A \mid E_i\right)$$

Proof: As A is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$

$\therefore\ A = A \cap (E_1 \cup E_2 \cup \ldots \cup E_n)$ [$\because$ if A is sub-set of B, then $A = A \cap B$] A

$\Rightarrow A = (A \cap E_1) \cup (A \cap E_2) \cup \ldots \cup (A \cap E_n)$ [Distributive property of set theory]

$= (E_1 \cap A) \cup (E_2 \cap A) \cup \ldots \cup (E_n \cap A)$

The above expression can be understood theoretically/logically also as explained under :

A happens in any of the following mutually exclusive ways:

(E_1 happens and then A happens) or (E_2 happens and then A happens) or (E_3 happens and then A happens) or … or (E_n happens and then A happens).
Now, as meanings of 'and' and 'or' in set theory are '$\cap$' and '$\cup$' respectively

$\therefore\ A = [(E_1 \cap A) \cup (E_2 \cap A) \cup (E_3 \cap A) \cup \ldots \cup (E_n \cap A)]$

$\Rightarrow P(A) = P[(E_1 \cap A) \cup (E_2 \cap A) \cup \ldots \cup (E_n \cap A)]$

$= P(E_1 \cap A) + P(E_2 \cap A) + \ldots + P(E_n \cap A)$

[$\because$ $E_1, E_2, \ldots, E_n$ and hence $E_1 \cap A, E_2 \cap A, \ldots, E_n \cap A$ are mutually exclusive]

$= P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2) + \ldots + (E_n)\, P(A|E_n)$

[Using multiplication theorem for dependent events]

$$= \sum_{i=1}^{n} P\left(E_i\right) P\left(A \mid E_i\right)$$

Hence proved

4.3 APPLICATIONS OF LAW OF TOTAL PROBABILITY

Here, in this section, we are going to take up various situations through examples, where the law of total probability is applicable.

Example 1: There are two bags. First bag contains 5 red, 6 white balls and the second bag contains 3 red, 4 white balls. One bag is selected at random and a ball is drawn from it. What is the probability that it is (i) red, (ii) white.

Solution: Let E_1 be the event that first bag is selected and E_2 be the event that second bag is selected.

$$\therefore \ P(E_1) = P(E_2) = \frac{1}{2}.$$

(i) Let R be the event of getting a red ball from the selected bag.

$$\therefore P(R \mid E_1) = \frac{5}{11}, \text{ and } P(R \mid E_2) = \frac{3}{7}.$$

Thus, the required probability is given by

$$P(R) = P(E_1)P(R \mid E_1) + P(E_2)P(R \mid E_2)$$

$$= \frac{1}{2} \times \frac{5}{11} + \frac{1}{2} \times \frac{3}{7}$$

$$= \frac{5}{22} + \frac{3}{14} = \frac{35+33}{154} = \frac{68}{154} = \frac{34}{77}$$

(ii) Let W be the event of getting a white ball from the selected bag.

$$\therefore P(W \mid E_1) = \frac{6}{11}, \text{ and } P(W \mid E_2) = \frac{4}{7}.$$

Thus, the required probability is given by

$$P(W) = P(E_1)P(W \mid E_1) + P(E_2)P(W \mid E_2)$$

$$= \frac{1}{2} \times \frac{6}{11} + \frac{1}{2} \times \frac{4}{7} = \frac{3}{11} + \frac{2}{7} = \frac{21+22}{77} = \frac{43}{77}.$$

Example 2: A factory produces certain type of output by 3 machines. The respective daily production figures are-machine X : 3000 units, machine Y: 2500 units and machine Z: 4500 units. Past experience shows that 1% of the output produced by machine X is defective. The corresponding fractions of defectives for the other two machines are 1.2 and 2 percent respectively. An item is drawn from the day's production. What is the probability that it is defective?

Solution: Let E_1, E_2 and E_3 be the events that the drawn item is produced by machine X, machine Y and machine Z, respectively. Let A be the event that the drawn item is defective.

As the total number of units produced by all the machines is

3000 + 2500 + 4500 = 10000,

$\therefore$ $P(E_1) = \dfrac{3000}{10000} = \dfrac{3}{10}$, $P(E_2) = \dfrac{2500}{10000} = \dfrac{1}{4}$, $P(E_3) = \dfrac{4500}{10000} = \dfrac{9}{20}$.

$P(A \mid E_1) = \dfrac{1}{100} = 0.01$, $P(A \mid E_2) = \dfrac{1.2}{100} = 0.012$, $P(A \mid E_3) = \dfrac{2}{100} = 0.02$.

Thus, the required probability = Probability that the drawn item is defective

$$= P(A)$$

$$= P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2) + P(E_3)\, P(A|E_3)$$

$$= \frac{3}{10} \times 0.01 + \frac{1}{4} \times 0.012 + \frac{9}{20} \times 0.02$$

$$= \frac{3}{1000} + \frac{3}{1000} + \frac{9}{1000}$$

$$= \frac{15}{1000} = 0.015.$$

Example 3: There are two coins-one unbiased and the other two- headed, otherwise they are identical. One of the coins is taken at random without seeing it and tossed. What is the probability of getting head?

Solution: Let E_1 and E_2 be the events of selecting the unbiased coin and the two-headed coin respectively. Let A be the event of getting head on the tossed coin.

$\therefore$ $P(E_1) = \dfrac{1}{2}$, $P(E_2) = \dfrac{1}{2}$ [$\because$ selection of each of the coin is equally likely]

$P(A \mid E_1) = \dfrac{1}{2}$ [$\because$ if it is unbiased coin, then head and tail are equally likely]

$P(A|E_2) = 1$ [$\because$ if it is two-headed coin, then getting the head is certain]

Thus, the required probability = P(A)

$$= P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2)$$

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

$$= \frac{1}{4} + \frac{1}{2} = \frac{3}{4}.$$

Example 4: The probabilities of selection of 3 persons for the post of a principal in a newly started colleage are in the ratio $4:3:2$. The probabilities that they will introduce co-education in the college are 0.2, 0.3 and 0.5, respectively. Find the probability that co-education is introduced in the college.

Solution: Let E_1, E_2, E_3 be the events of selection of first, second and third person for the post of a principal respectively. Let A be the event that co-education is introduced.

$\therefore$ $P(E_1) = \dfrac{4}{9}$, $P(E_2) = \dfrac{3}{9}$, $P(E_3) = \dfrac{2}{9}$

$P(A \mid E_1) = 0.2$, $P(A \mid E_2) = 0.3$, $P(A \mid E_3) = 0.5$.

Thus, the required probability = P(A)

$$= P(E_1)P(A|E_1) + P(E_2)P(A|E_2) + P(E_3)\,P(A|E_3)$$

$$= \frac{4}{9}\times 0.2 + \frac{3}{9}\times 0.3 + \frac{2}{9}\times 0.5$$

$$= \frac{0.8}{9} + \frac{0.9}{9} + \frac{1}{9} = \frac{2.7}{9} = 0.3$$

Now, you can try the following exercises.

E1) A person gets a construction job and agrees to undertake it. The completion of the job in time depends on whether there happens to be strike or not in the company. There are 40% chances that there will be a strike. Probability that job is completed in time is 30% if the strike takes place and is 70% if the strike does not take place. What is the probability that the job will be completed in time?

E2) What is the probability that a year selected at random will contains 53 Sundays?

E3) There are two bags, first bag contains 3 red, 5 black balls and the second bag contains 4 red, 5 black balls. One ball is drawn from the first bag and is put into the second bag without noticing its colour. Then two balls are drawn from the second bag. What is the probability that balls are of opposite colours?

4.4 BAYES' THEOREM

In Sec. 4.2 of this unit, we have discussed that if we are interested in finding the probability of the event of second stage, then it is obtained using law of total probability. But if the happening of the event of second stage is given to us and on this basis we find the probability of the events of first stage, then the probability of an event of first stage is the revised (or posterior) probabilities and is obtained using an important theorem known as Bayes' theorem given by Thomas Bayes (died in 1761, at the age of 59), a British Mathematician, published after his death in 1763. This theorem is also known as 'Inverse probability theorem', because here moving from first stage to second stage, we again find the probabilities (revised) of the events of first stage i.e. we move inversely. Thus, using this theorem, probabilities can be revised on the basis of having some related new information.

As an illustration, let us consider the same example as taken in Sec. 4.2 of this unit. In this example, if we are given that the drawn ball is of particular colour and it is asked to find, on this basis, the probability that Urn I or Urn II was selected, then these are the revised (posterior) probabilities and are obtained using Bayes' theorem, which is stated and proved as under:

Statement: Let S be the sample space and $E_1, E_2, \ldots, E_n$ be n mutually exclusive and exhaustive events with $P(E_i) \neq 0$; i = 1, 2, .., n. Let A be any event which is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$ (i.e. at least one of the events $E_1, E_2, \ldots, E_n$) with P(A) > 0 [Notice that up to this line the statement is same as that of law of total probability], then

$$P(E_i \mid A) = \frac{P(E_i)P(A \mid E_i)}{P(A)}, i = 1, 2, ..., n$$

where P(A) = P(E_1) P(A|E_1) + P(E_2) P(A|E_2) + … +P(E_n) P(A|E_n).

Proof: First you have to prove that

P(A) = P(E_1) P(A|E_1) + P(E_2) P(A|E_2) +…+P(E_n) P(A|E_n).

which is nothing but the law of total probability that has already been proved in Sec. 4.2. After proving this, proceed as under:

$$P(E_i \mid A) = \frac{P(E_i \cap A)}{P(A)} \qquad \left[\begin{array}{l}\text{Applying the formula of conditional} \\ \text{probability i.e. } P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}\end{array}\right]$$

$$= \frac{P(E_i)P(A \mid E_i)}{P(A)} \quad \left[\begin{array}{l}\text{Applying multiplication theorem for dependent} \\ \text{events i.e. } P(A \cap B) = P(A)P(B \mid A)\end{array}\right]$$

4.5 APPLICATIONS OF BAYES' THEOREM

Example 5: Let us consider the problem given in Example 1 of Sec 4.3 of this unit after replacing the question asked therein (i.e. the last sentence) by the following question:

If it is found to be red, what is the probability of?

i) selecting the first bag

ii) selecting the second bag

Solution: First, we have to give the solution exactly as given for Example 1 of Sec. 4.3 of this unit. After that, we are to proceed as follows:

i) Probability of selecting the first bag given that the ball drawn is red

$$= P(E_1 \mid R)$$

$$= \frac{P(E_1)P(R \mid E_1)}{P(R)} \qquad \left[\begin{array}{l}\text{Applying Bayes'} \\ \text{theorem}\end{array}\right]$$

$$= \frac{\frac{1}{2} \times \frac{5}{11}}{\frac{34}{77}} = \frac{5}{22} \times \frac{77}{34} = \frac{35}{68}$$

ii) Probability of selecting the second bag given that the ball drawn is red

$$P(E_2|R) = \frac{P(E_2)P(R \mid E_2)}{P(R)} = \frac{\frac{1}{2} \times \frac{3}{7}}{\frac{34}{77}} = \frac{3}{14} \times \frac{77}{34} = \frac{33}{68}$$

Example 6: Consider the problem given in Example 2 of Sec. 4.3 of this unit after replacing the question asked therein (i.e. the last sentence) by the following question:

If the drawn item is found to be defective, what is the probability that it has been produced by machine Y?

Solution: Proceed exactly in the manner the Example 2 of Sec. 4.3 of this unit has been solved and then as under:

Probability that the drawn item has been produced by machine Y given that it is defective

$$= P(E_2 \mid A)$$

$$= \frac{P(E_2)P(A \mid E_2)}{P(A)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

$$= \frac{\frac{1}{4} \times 0.012}{0.015} = \frac{0.003}{0.015} = \frac{1}{5}.$$

Example 7: Consider the problem given in Example 3 of Sec. 4.3 of this unit after replacing the question asked therein (i.e. the last sentence) by the following question:

If head turns up, what is the probability that

i) it is the two-headed coin

ii) it is the unbiased coin.

Solution: First give the solution of Example 3 of Sec. 4.3 of this unit and then proceed as under:

Now, the probability that the tossed coin is two-headed given that head turned up

$$= P(E_2 \mid A)$$

$$= \frac{P(E_2)P(A \mid E_2)}{P(A)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

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

ii) The probability that the tossed coin is unbiased given that head turned up

$$= P(E_1 \mid A)$$

$$= \frac{P(E_1)P(A \mid E_1)}{P(A)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

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

Example 8: Consider the statement given in Example 4 of Sec. 4.3 of this unit after replacing the question asked therein (i.e. the last sentence) by the following question:

If the co-education is introduced by the candidate selected for the post of principal, what is the probability that first candidate was selected.

Solution: First give the solution of Example 4 of Sec. 4.3 of this unit and then proceed as under:

The required probability = $P\left(E_1 \mid A\right)$

$$= \frac{P\left(E_1\right)P\left(A \mid E_1\right)}{P\left(A\right)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

$$= \frac{\frac{4}{9}\times 0.2}{0.3} = \frac{4}{9}\times\frac{0.2}{0.3} = \frac{8}{27}.$$

Now, you can try the following exercises.

E4) A bag contains 4 red and 5 white balls. Another bag contains 2 red and 3 white balls. A ball is drawn from the first bag and is transferred to the second bag. A ball is then drawn from the second bag and is found to be red, what is the probability that red ball was transferred from first to second bag?

E5) An insurance company insured 1000 scooter drivers, 3000 car drivers and 6000 truck drivers. The probabilities that scooter, car and truck drivers meet an accident are 0.02, 0.04, 0.25 respectively. One of the insured persons meets with an accident. What is the probability that he is a
(i) car driver

(ii) truck driver

E6) By examining the chest X-ray, the probability that T.B is detected when a person is actually suffering from T.B. is 0.99. The probability that the doctor diagnoses incorrectly that a person has T.B. on the basis of X-ray is 0.002. In a certain city, one in 1000 persons suffers from T.B. A person is selected at random and is diagnosed to have T.B., what is the chance that he actually has T.B.?

E7) A person speaks truth 3 out of 4 times. A die is thrown. She reports that there is five. What is the chance there was five?

4.6 SUMMARY

In this unit, we have covered the following points:

1) There are experiments which are conducted in two stages for completion. Such experiments are termed as two-stage experiments. At the first stage, the experiment involves selection of one of the given number of possible mutually exclusive events. At the second stage, the experiment involves happening of an event which is a sub-set of at least one of the events of first stage. If we are interested in finding the probability of the event of second stage, then it is obtained using **law of total probability**, but if the

happening of the event of the second stage is given to us and on this basis we find the probability of an event of the first stage, then this probability of the event of the first stage is the revised (or posterior) probability and is obtained by using **Bayes' theorem.**

2) **Law of total probability** Let S be the sample space and $E_1, E_2, \ldots, E_n$ be n mutually exclusive and exhaustive events with $P(E_i) \neq 0$; i = 1, 2, …, n. Let A be any event which is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$ (i.e. at least one of the events $E_1, E_2, \ldots, E_n$) with P(A) > 0, then

$$P(A) = P(E_1)\,P(A|E_1) + P(E_2)\,P(A|E_2) + \ldots + P(E_n)\,P(A|E_n)$$

$$= \sum_{i=1}^{n} P(E_i)\,P(A\,|\,E_i).$$

3) **Bayes' theorem** is an extension of law of total probability and is stated as:

Let S be the sample space. Let $E_1, E_2, \ldots, E_n$ be n mutually exclusive and exhaustive events with $P(E_i) \neq 0$; i = 1, 2, …, n. Let A be any event which is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$ (i.e. at least one of the events $E_1, E_2, \ldots, E_n$) with P(A) > 0 [Notice that up to this line the statement is as same as that of law of total probability], then

$$P(E_i\,|\,A) = \frac{P(E_i)\,P(A\,|\,E_i)}{P(A)}, i = 1, 2, \ldots, n$$

where $P(A) = P(E_1)\,P(A|E_1) + P(E_2)\,P(A|E_2) + \ldots + P(E_n)\,P(A|E_n)$.

4.7 SOLUTIONS/ANSWERS

E1) Let E_1 be the event that strike takes place and E_2 be the event that strike does not take place.

Let A be the event that the job will be completed in time.

$$\therefore P(E_1) = 40\% = \frac{40}{100}, P(E_2) = 1 - P(E_1) = 1 - \frac{40}{100} = \frac{60}{100},$$

$$P(A\,|\,E_1) = 30\% = \frac{30}{100},\ P(A\,|\,E_2) = 70\% = \frac{70}{100}$$

Thus, the required probability is

$$P(A) = P(E_1)P(A\,|\,E_1) + P(E_2)P(A\,|\,E_2)$$

$$= \frac{40}{100} \times \frac{30}{100} + \frac{60}{100} \times \frac{70}{100}$$

$$= \frac{1200}{10000} + \frac{4200}{10000} = \frac{12}{100} + \frac{42}{100} = \frac{54}{100} = 54\%$$

E2) Such probability in case of a leap year and in case of a non-leap year has already been obtained in Unit 1 of this course. But here we do not know as to whether it is a leap year or non-leap year, hence we will have to do it using law of total probability considering both the possibilities i.e. we shall proceed as follows:

Let E_1 be the event that the year selected is a leap year and E_2 be the event that it is a non-leap year. Let A be the event that the selected year contains 53 Sundays.

We know that out of 4 consecutive years, 1 is a leap year and 3 are non-leap years, therefore,

$$P(E_1) = \frac{1}{4},\ P(E_2) = \frac{3}{4}.$$

A leap years consists of 366 days, i.e. 52 complete weeks and 2 over days. These 2 over days may be "Sunday and Monday", "Monday and Tuesday", "Tuesday and Wednesday" ," Wednesday and Thursday", "Thursday and Friday", " Friday and Saturday" and " Saturday and Sunday". Out of these 7 cases, Sunday is included in 2 cases and hence the probability that a leap year will consist of 53 Sundays is

$$P(A \mid E_1) = \frac{2}{7}.$$

A non-leap year consists of 365 days, i.e. 52 complete weeks and 1 over day. This over day may be one of the seven days of the week and hence the probability that a non-leap year will consist of 53 sundays is

$$P(A \mid E_2) = \frac{1}{7}.$$

$\therefore$ By the total law of probability, the probability that a year selected at random will consists of 53 Sundays is

$$P(A) = P(E_1)P(A \mid E_1) + P(E_2)P(A \mid E_2)$$

$$= \frac{1}{4}\times\frac{2}{7}+\frac{3}{4}\times\frac{1}{7} = \frac{2}{28}+\frac{3}{28}=\frac{5}{28}.$$

E3) We are given that first of all one ball is transferred from bag I to bag II and then two balls are drawn from bag II. This job can be done in two mutually exclusive ways

(a) A red ball is transferred from bag I to bag II and then two balls are drawn from bag II

(b) A black ball is transferred from bag I to bag II and then two balls are drawn from bag II

We define the following events:

E_1 be the event of getting a red ball from bag I

E_2 be the event of getting a black ball from bag I

A be the event of getting two balls of opposite colours from the bag II

$$\therefore P(E_1) = \frac{{}^{3}C_1}{{}^{8}C_1} = \frac{3}{8}$$

$$P(E_2) = \frac{{}^{5}C_1}{{}^{8}C_1} = \frac{5}{8}$$

$$P(A|E_1) = \frac{{}^5C_1 \times {}^5C_1}{{}^{10}C_2} = \frac{5\times5}{45} = \frac{5}{9}.$$

$$P(A|E_2) = \frac{{}^4C_1 \times {}^6C_1}{{}^{10}C_2} = \frac{4\times6}{45} = \frac{8}{15}$$

Hence, the required probability is given by

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2)$$

$$= \frac{3}{8}\times\frac{5}{9} + \frac{5}{8}\times\frac{8}{15}$$

$$= \frac{5}{24} + \frac{1}{3} = \frac{5+8}{24} = \frac{13}{24}$$

E4) Let E_1 be the event that a red ball is drawn from the first bag and E_2 be the event that the drawn ball from the first bag is white. Let A be the event of drawing a red ball from the second bag after transferring the ball drawn from first bag into it.

$$\therefore\ P(E_1) = \frac{4}{9}, P(E_2) = \frac{5}{9}$$

$$P(A|E_1) = \frac{3}{6}, P(A|E_2) = \frac{2}{6}$$

$\therefore$ By law of total probability,

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2)$$

$$= \frac{4}{9}\times\frac{3}{6} + \frac{5}{9}\times\frac{2}{6} = \frac{11}{27}.$$

Thus, the required probability = $P(E_1|A)$

$$= \frac{P(E_1)P(A|E_1)}{P(A)} \quad \left[\begin{array}{l}\text{Applying Bayes'}\\ \text{theorem}\end{array}\right]$$

$$= \frac{\frac{4}{9}\times\frac{3}{6}}{\frac{11}{27}} = \frac{6}{11}$$

E5) Total number of drivers = 1000 + 3000 + 6000 = 10000.

Let E_1, E_2, E_3 be the events that the selected insured person is a scooter driver, car driver, truck driver respectively. Let A be the event that a driver meets with an accident.

$$\therefore\ P(E_1) = \frac{1000}{10000} = \frac{1}{10},\ P(E_2) = \frac{2000}{10000} = \frac{1}{5},\ P(E_3) = \frac{6000}{10000} = \frac{6}{10} = \frac{3}{5}.$$

$$P(A|E_1) = 0.02 = \frac{2}{100},\ P(A|E_2) = 0.04 = \frac{4}{100},\ P(A|E_3) = 0.25 = \frac{25}{100}$$

$\therefore$ By total probability theorem, we have

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2) + P(E_3)P(A|E_3)$$

$$= \frac{1}{10}\times\frac{2}{100}+\frac{1}{5}\times\frac{4}{100}+\frac{3}{5}\times\frac{25}{100}$$

$$= \frac{1}{500}+\frac{4}{500}+\frac{75}{500}=\frac{80}{500}=\frac{4}{25}$$

i) The required probability $= P(E_2 | A)$

$$= \frac{P(E_2)P(A|E_2)}{P(A)} \quad \left[\begin{array}{l}\text{Applying Bayes'}\\ \text{theorem}\end{array}\right]$$

$$= \frac{\frac{1}{5}\times\frac{4}{100}}{\frac{4}{25}} = \frac{4}{500}\times\frac{25}{4}=\frac{1}{20}.$$

ii) The required probability $= P(E_3 | A)$

$$= \frac{P(E_3)P(A|E_3)}{P(A)} \quad \left[\begin{array}{l}\text{Applying Bayes'}\\ \text{theorem}\end{array}\right]$$

$$= \frac{\frac{3}{5}\times\frac{25}{100}}{\frac{4}{25}} = \frac{75}{500}\times\frac{25}{4}=\frac{75}{80}=\frac{15}{16}.$$

E6) Let E_1 be the event that a person selected from the population of the city is suffering from the T.B. and E_2 be the event that he/she is not suffering from the T.B. Let A be the event that the selected person is diagnosed to have T.B. Therefore, according to given,

$$P(E_1) = \frac{1}{1000} = 0.001,\quad P(E_2) = 1 - 0.001 = 0.999 = \frac{999}{1000},$$

$$P(A|E_1) = 0.99 = \frac{99}{100}, P(A|E_2) = 0.002 = \frac{2}{1000}.$$

$\therefore$ By total probabilities theorem, we have

$$P(A) = P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2)$$

$$= \frac{1}{1000}\times\frac{99}{100}+\frac{999}{1000}\times\frac{2}{1000}$$

$$= \frac{99}{100000}+\frac{1998}{1000000} = \frac{990+1998}{1000000}=\frac{2988}{1000000}$$

Thus, the required probability = $P(E_1|A)$

$$= \frac{P(E_1)P(A|E_1)}{P(A)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

$$= \frac{\frac{1}{1000} \times \frac{99}{100}}{\frac{2988}{1000000}} = \frac{990}{2988} = \frac{55}{166}.$$

E7) Let E_1 be the event that the person speaks truth, E_2 be the event that she tells a lie and A be the event that she reports a five.

$$\therefore\ P(E_1) = \frac{3}{4},\ P(E_2) = \frac{1}{4},\ P(A|E_1) = \frac{1}{6},\ P(A|E_2) = \frac{5}{6}.$$

By law of total probability, we have

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2)$$

$$= \frac{3}{4} \times \frac{1}{6} + \frac{1}{4} \times \frac{5}{6}$$

$$= \frac{3}{24} + \frac{5}{24} = \frac{8}{24} = \frac{1}{3}.$$

Thus, the required probability = $P(E_1|A)$

$$= \frac{P(E_1)P(A|E_1)}{P(A)} \quad \begin{bmatrix}\text{Applying Bayes'} \\ \text{theorem}\end{bmatrix}$$

$$= \frac{\frac{3}{4} \times \frac{1}{6}}{\frac{1}{3}} = \frac{3}{4} \times \frac{1}{6} \times \frac{3}{1} = \frac{3}{8}.$$

UNIT 5 RANDOM VARIABLES

Structure

5.1 INTRODUCTION

In the previous units, we have studied the assignment and computation of probabilities of events in detail. In those units, we were interested in knowing the occurrence of outcomes. In the present unit, we will be interested in the numbers associated with such outcomes of the random experiments. Such an interest leads to study the concept of random variable.

In this unit, we will introduce the concept of random variable, discrete and continuous random variables in Sec. 5.2 and their probability functions in Secs. 5.3 and 5.4.

Objectives

A study of this unit would enable you to:

- define a random variable, discrete and continuous random variables;
- specify the probability mass function, i.e. probability distribution of discrete random variable;
- specify the probability density function, i.e. probability function of continuous random variable; and
- define the distribution function.

5.2 RANDOM VARIABLE

Study related to performing the random experiments and computation of probabilities for events (subsets of sample space) have been made in detail in the first four units of this course. In many experiments, we may be interested in a numerical characteristic associated with outcomes of a random experiment. Like the outcome, the value of such a numerical characteristic cannot be predicted in advance.

For example, suppose a die is tossed twice and we are interested in number of times an odd number appears. Let X be the number of appearances of odd number. If a die is thrown twice, an odd number may appear ‘0’ times (i.e. we may have even number both the times) or once (i.e. we may have odd number in one throw and even number in the other throw) or twice (i.e. we may have odd number both the times). Here, X can take the values 0, 1, 2 and is a variable quantity behaving randomly and hence we may call it as 'random variable'. Also notice that its values are real and are defined on the sample space

{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 1), (5, 2), (5, 3), (5, 4), (5, 4), (5, 6),
(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

i.e.

$$X = \begin{cases} 0, \text{ if the outcome is } (2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2), \\ \qquad\qquad\qquad (6,4),(6,6) \\ 1, \text{ if the outcome is } (1,2),(1,4),(1,6),(2,1),(2,3),(2,5),(3,2), \\ \qquad\qquad\qquad (3,4),(3,6),(4,1),(4,3),(4,5),(5,2),(5,4), \\ \qquad\qquad\qquad (5,6),(6,1),(6,3),(6,5) \\ 2, \text{ if the outcome is } (1,1),(1,3),(1,5),(3,1),(3,3),(3,5),(5,1), \\ \qquad\qquad\qquad (5,3),(5,5) \end{cases}$$

So, $P[X = 0] = \frac{9}{36} = \frac{1}{4}, P[X = 1] = \frac{18}{36} = \frac{1}{2}, P[X = 2] = \frac{9}{36} = \frac{1}{4}$,

and $P[X = 0] + P[X = 1] + P[X = 2] = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$

Observe that, a probability can be assigned to the event that X assumes a particular value. It can also be observed that the sum of the probabilities corresponding to different values of X is one.

So, a random variable can be defined as below:

Definition: A random variable is a real-valued function whose domain is a set of possible outcomes of a random experiment and range is a sub-set of the set of real numbers and has the following properties:

i) Each particular value of the random variable can be assigned some probability

ii) Uniting all the probabilities associated with all the different values of the random variable gives the value 1(unity).

Remark 1: We shall denote random variables by capital letters like X, Y, Z, etc. and write r.v. for random variable.

5.3 DISCRETE RANDOM VARIABLE AND PROBABILITY MASS FUNCTION

Discrete Random Variable

A random variable is said to be discrete if it has either a finite or a countable number of values. Countable number of values means the values which can be arranged in a sequence, i.e. the values which have one-to-one correspondence with the set of natural numbers, i.e. on the basis of three-four successive known terms, we can catch a rule and hence can write the subsequent terms. For example suppose X is a random variable taking the values say 2, 5, 8, 11, … then we can write the fifth, sixth, … values, because the values have one-to-one correspondence with the set of natural numbers and have the general term as $3n - 1$ i.e. on taking n = 1, 2, 3, 4, 5, … we have 2, 5, 8, 11, 14,…. So, X in this example is a discrete random variable. The number of students present each day in a class during an academic session is an example of discrete random variable as the number cannot take a fractional value.

Probability Mass Function

Let X be a r.v. which takes the values x_1, x_2, ... and let $P[X = x_i] = p(x_i)$. This function $p(x_i)$, i =1,2, … defined for the values x_1, x_2, … assumed by X is called probability mass function of X satisfying $p(x_i) \geq 0$ and $\sum_i p(x_i) = 1$.

The set $\{(x_1, p(x_1)), (x_2, p(x_2)), ...\}$ specifies the probability distribution of a discrete r.v. X. Probability distribution of r.v. X can also be exhibited in the following manner:

X	x_1	x_2	$x_3 \ldots$
$p(x)$	$p(x_1)$	$p(x_2)$	$p(x_3) \ldots$

Now, let us take up some examples concerning probability mass function:

Example 1: State, giving reasons, which of the following are not probability distributions:

(i)

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

(ii)

X	0	1	2
$p(x)$	$\frac{3}{4}$	$-\frac{1}{2}$	$\frac{3}{4}$

(iii)

X	0	1	2
p(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(iv)

X	0	1	2	3
p(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{8}$

Solution:

(i) Here $p(x_i) \geq 0$, i = 1, 2; but

$$\sum_{i=1}^{2} p(x_i) = p(x_1) + p(x_2) = p(0) + p(1) = \frac{1}{2} + \frac{3}{4} = \frac{5}{4} > 1.$$

So, the given distribution is not a probability distribution as $\sum_{i=1}^{2} p(x_i)$ is greater than 1.

(ii) It is not probability distribution as $p(x_2) = p(1) = -\frac{1}{2}$ i.e. negative

(iii) Here, $p(x_i) \geq 0$, i = 1, 2, 3

and $\sum_{i=1}^{3} p(x_i) = p(x_1) + p(x_2) + p(x_3) = p(0) + p(1) + p(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$

$\therefore$ The given distribution is probability distribution.

(iv) Here, $p(x_i) \geq 0$, i = 1, 2, 3, 4; but

$$\sum_{i=1}^{4} p(x_i) = p(x_1) + p(x_2) + p(x_3) + p(x_4)$$

$$= p(0) + p(1) + p(2) + p(3) = \frac{1}{8} + \frac{3}{8} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} < 1.$$

$\therefore$ The given distribution is not probability distribution.

Example 2: For the following probability distribution of a discrete r.v. X, find

i) the constant c,

ii) $P[X \leq 3]$ and

iii) $P[1 < X < 4]$.

X	0	1	2	3	4	5
$p(x)$	0	c	c	2c	3c	c

Solution:

i) As the given distribution is probability distribution,

$$\therefore \quad \sum_i p(x_i) = 1$$

$$\Rightarrow 0 + c + c + 2c + 3c + c = 1 \Rightarrow 8\,c = 1 \Rightarrow \; c = \frac{1}{8}$$

ii) $P[X \le 3] = P[X = 3] + P[X = 2] + P[X = 1] + P[X = 0]$

$$= 2c + c + c + 0 = 4\,c = 4 \times \frac{1}{8} = \frac{1}{2}.$$

iii) $P[1 < X < 4] = P[X = 2] + P[X = 3] = c + 2c = 3c = 3 \times \frac{1}{8} = \frac{3}{8}.$

Example 3: Find the probability distribution of the number of heads when three fair coins are tossed simultaneously.

Solution: Let X be the number of heads in the toss of three fair coins.

As the random variable, "the number of heads" in a toss of three coins may be 0 or 1 or 2 or 3 associated with the sample space

{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT},

$\therefore$ X can take the values 0, 1, 2, 3, with

$P[X = 0] = P[TTT\,] = \frac{1}{8}$

$P[X = 1] = P[HTT, THT, TTH] = \frac{3}{8}$

$P[X = 2] = P[HHT, HTH, THH] = \frac{3}{8}$

$P[X = 3] = P\,[HHH] = \frac{1}{8}.$

$\therefore$ Probability distribution of X, i.e. the number of heads when three coins are tossed simultaneously is

X	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

which is the required probability distribution.

Example 4: A r.v. X assumes the values $-2, -1, 0, 1, 2$ such that

$P[X = -2] = P[X = -1] = P[X = 1] = P[X = 2],$

$P[X < 0] = P[X = 0] = P[X > 0].$

Obtain the probability mass function of X.

Solution: As P[X < 0] = P[X = 0] = P[X > 0]

$\therefore$ P[X = − 1] + P[X = −2] = P[X = 0] = P[X = 1] + P[X = 2]

$\Rightarrow$ p + p = P[X = 0] = p + p

[Letting P[X = 1] = P[X = 2] = P[X = −1] = P[X = −2] = p]

$\Rightarrow$ P[X = 0] = 2p.

Now, as P[X < 0] + P[X = 0] + P[X > 0] = 1,

$\therefore$ P[X = −1] + P[X = −2] + P[X = 0] + P[X = 1] + P[X = 2] = 1

$\Rightarrow$ p + p + 2p + p + p = 1

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

$\therefore P[X = 0] = 2p = 2 \times \frac{1}{6} = \frac{2}{6}$,

$P[X = -1] = P[X = -2] = P[X = 1] = P[X = 2] = p = \frac{1}{6}$.

Hence, the probability distribution of X is given by

X	−2	− 1	0	1	2
p(x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Now, here are some exercises for you.

E1) 2 bad articles are mixed with 5 good ones. Find the probability distribution of the number of bad articles, if 2 articles are drawn at random.

E2) Given the probability distribution:

X	0	1	2	3
p(x)	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{10}$

Let $Y = X^2 + 2X$. Find the probability distribution of Y.

E3) An urn contains 3 white and 4 red balls. 3 balls are drawn one by one with replacement. Find the probability distribution of the number of red balls.

Let us define and explain a continuous random variable and its probability function in the next section.

5.4 CONTINUOUS RANDOM VARIABLE AND PROBABILITY DENSITY FUNCTION

In Sec. 5.3 of this unit, we have defined the discrete random variable as a random variable having countable number of values, i.e. whose values can be arranged in a sequence. But, if a random variable is such that its values cannot be arranged in a sequence, it is called continuous random variable. Temperature of a city at various points of time during a day is an example of continuous random variable as the temperature takes uncountable values, i.e. it can take fractional values also. So, a random variable is said to be continuous if it can take all possible real (i.e. integer as well as fractional) values between two certain limits. For example, let us denote the variable, "Difference between the rainfall (in cm) of a city and that of another city on every rainy day in a rainy reason", by X, then X here is a continuous random variable as it can take any real value between two certain limits. It can be noticed that for a continuous random variable, the chance of occurrence of a particular value of the variable is very small, so instead of specifying the probability of taking a particular value by the variable, we specify the probability of its lying within an interval. For example, chance that an athlete will finish a race in say exactly 10 seconds is very-very small, i.e. almost zero as it is very rare to finish the race in a fixed time. Here, the probability is specified for an interval, i.e. we may be interested in finding as to what is the probability of finishing the race by the athlete in an interval of say 10 to 12 seconds.

So, continuous random variable is represented by different representation known as **probability density function** unlike the discrete random variable which is represented by probability mass function.

Probability Density Function

Let $f(x)$ be a continuous function of x. Suppose the shaded region ABCD shown in the following figure represents the area bounded by $y = f(x)$, x-axis and the ordinates at the points x and $x + \delta x$, where δx is the length of the interval $(x, x + \delta x)$.

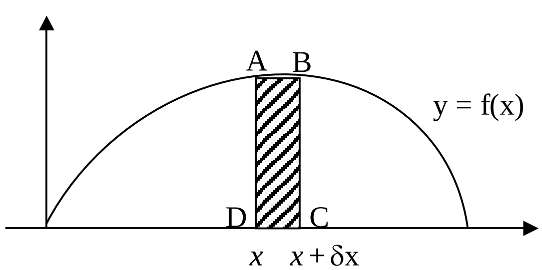


Fig. 5.1

Now, if δx is very-very small, then the curve AB will act as a line and hence the shaded region will be a rectangle whose area will be AD × DC i.e. $f(x)\,\delta x$ [$\because$ AD = the value of y at x i.e. $f(x)$, DC = length δx of the interval $(x, x+\delta x)$]

Also, this area = probability that X lies in the interval $(x, x+\delta x)$

$$= P[x \le X \le x+\delta x]$$

Hence,

$$P[x \le X \le x + \delta x] = f(x)\,\delta x$$

$$\Rightarrow \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x), \text{ where } \delta x \text{ is very-very small}$$

$$\Rightarrow \lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x).$$

f(x), so defined, is called probability density function.

Probability density function has the same properties as that of probability mass function. So, $f(x) \ge 0$ and sum of the probabilities of all possible values that the random variable can take, has to be 1. But, here, as X is a continuous random variable, the summation is made possible through 'integration' and hence

$$\int_R f(x)\,dx = 1,$$

where integral has been taken over the entire range R of values of X.

Remark 2

i) Summation and integration have the same meanings but in mathematics there is still difference between the two and that is that the former is used in case of discrete values, i.e. countable values and the latter is used in continuous case.

ii) An essential property of a continuous random variable is that there is zero probability that it takes any specified numerical value, but the probability that it takes a value in specified intervals is non-zero and is calculable as a definite integral of the probability density function of the random variable and hence the probability that a continuous r.v. X will lie between two values a and b is given by

$$P[a < X < b] = \int_a^b f(x)\,dx.$$

Example 5: A continuous random variable X has the probability density function:

$f(x) = Ax^3,\ 0 \le x \le 1.$

Determine

i) A

ii) $P[0.2 < X < 0.5]$

iii) $P\left[X > \frac{3}{4} \text{ given } X > \frac{1}{2}\right]$

Solution:

(i) As f(x) is probability density function,

$$\therefore \int_R f(x)\,dx = 1$$

$$\Rightarrow \int_0^1 f(x)\,dx = 1 \Rightarrow \int_0^1 Ax^3 dx = 1$$

$$\Rightarrow A\left[\frac{x^4}{4}\right]_0^1 = 1 \Rightarrow A\left(\frac{1}{4} - 0\right) = 1 \Rightarrow A = 4$$

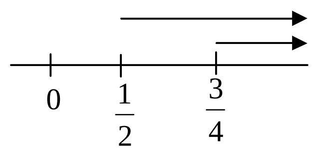
(ii) $P[0.2 < X < 0.5] = \int_{0.2}^{0.5} f(x)\,dx = \int_{0.2}^{0.5} Ax^3 dx = 4\left[\frac{x^4}{4}\right]_{0.2}^{0.5} = [(0.5)^4 - (0.2)^4]$

$= 0.0625 - 0.0016 = 0.0609$

(iii) $P[X > \frac{3}{4}$ given $X > \frac{1}{2}] = P[X > \frac{3}{4} \mid X > \frac{1}{2}]$

$$= \frac{P\left[X > \frac{3}{4} \cap X > \frac{1}{2}\right]}{P\left[X > \frac{1}{2}\right]} \qquad \left[\because P(A|B) = \frac{P(A \cap B)}{P(B)}\right]$$

$$= \frac{P\left[X > \frac{3}{4}\right]}{P\left[X > \frac{1}{2}\right]} \qquad \left[\begin{array}{l}\because \text{the common portion for} \\ X > \frac{3}{4} \text{ and } X > \frac{1}{2} \text{ is } X > \frac{3}{4}\end{array}\right]$$



Now, $P\left[X > \frac{3}{4}\right] = \int_{3/4}^{1} f(x)\,dx$

$$\left[\begin{array}{l}\because \text{ lower limit is } \frac{3}{4} \text{ and upper limit} \\ \text{is given in the problem which is } 1\end{array}\right]$$

$$= \int_{3/4}^{1} 4x^3 dx = 4\left[\frac{x^4}{4}\right]_{3/4}^{1} = (1)^4 - \left(\frac{3}{4}\right)^4 = 1 - \frac{81}{256} = \frac{175}{256}, \text{ and}$$

$$P\left[X > \frac{1}{2}\right] = \int_{1/2}^{1} f(x) = [x^4]_{1/2}^{1} = 1 - \frac{1}{16} = \frac{15}{16}.$$

$$\therefore \text{ the required probability } = \frac{P\left[X > \frac{3}{4}\right]}{P\left[X > \frac{1}{2}\right]} = \frac{175}{256} \times \frac{16}{15} = \frac{35}{16 \times 3} = \frac{35}{48}.$$

Example 6: The p.d.f. of the different weights of a "1 litre pure ghee pack" of a company is given by:

$$f(x) = \begin{cases} 200(x-1) & \text{for } 1 \le x \le 1.1 \\ 0, & \text{otherwise} \end{cases}$$

Examine whether the given p.d.f. is a valid one. If yes, find the probability that the weight of any pack will lie between 1.01 and 1.02.

Solution: For $1 \le x \le 1.1$, we have $f(x) \ge 0$, and

$$\int_1^{1.1} f(x)\,dx = \int_1^{1.1} 200(x-1)\,dx = 200\left[\frac{x^2}{2}-x\right]_1^{1.1} = 200\left[\left\{\frac{(1.1)^2}{2}-1.1\right\}-\left\{\frac{1}{2}-1\right\}\right]$$

$$= 200\left[\left(\frac{1.21-2.2}{2}\right)-\left(\frac{1-2}{2}\right)\right] = 200\left[-\frac{0.99}{2}+\frac{1}{2}\right] = 200\frac{(0.01)}{2} = 1.$$

$\therefore$ f(x) is p.d.f.

$$\text{Now, } P[1.01 < X < 1.02] = \int_{1.01}^{1.02} 200(x-1)\,dx = 200\left[\frac{x^2}{2}-x\right]_{1.01}^{1.02}$$

$$= 200\left[\left\{\frac{(1.02)^2}{2}-1.02\right\}-\left\{\frac{(1.01)^2}{2}-1.01\right\}\right]$$

$$= 200\left[\frac{1.0404}{2}-1.02-\frac{1.0201}{2}+1.01\right]$$

$$= 200\left[0.5202-1.02-0.51005+1.01\right]$$

$$= 200\left[0.00015\right] = 0.03.$$

Now, you can try the following exercise.

E4) The life (in hours) X of a certain type of light bulb may be supposed to be a continuous random variable with p.d.f.:

$$f(x) = \begin{cases} \dfrac{A}{x^3}, & 1500 < x < 2500 \\ 0, & \text{elsewhere} \end{cases}$$

Determine the constant A and compute the probability that $1600 \le X \le 2000$.

5.5 DISTRIBUTION FUNCTION

A function F defined for all values of a random variable X by $F(x) = P[X \le x]$ is called the distribution function. It is also known as the cumulative distribution function (c.d.f.) of X since it is the cumulative probability of X up to and including the value x. As X can take any real value, therefore the domain of the distribution function is set of real numbers and as F(x) is a probability value, therefore the range of the distribution function is [0, 1].

Remark 3: Here, X denotes the random variable and x represents a particular value of random variable. $F(x)$ may also be written as $F_X(x)$, which means that it is a distribution function of random variable X.

Discrete Distribution Function

Distribution function of a discrete random variable is said to be discrete distribution function or cumulative distribution function (c.d.f.). Let X be a discrete random variable taking the values x_1, x_2, x_3, … with respective probabilities p_1, p_2, p_3, …

Then $F(x_i) = P[X \le x_i] = P[X = x_1] + P[X = x_2] + \ldots + P[X = x_i]$

$$= p_1 + p_2 + \ldots + p_i.$$

The distribution function of X, in this case, is given as in the following table:

X	F(x)
x_1	p_1
x_2	$p_1 + p_2$
x_3	$p_1 + p_2 + p_3$
x_4	$p_1 + p_2 + p_3 + p_4$
.	.
.	.
.	.

The value of F(x) corresponding to the last value of the random variable X is always 1, as it is the sum of all the probabilities. F(x) remains 1 beyond this last value of X also, as it being a probability can never exceed one.

For example, Let X be a random variable having the following probability distribution:

X	0	1	2
p(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Notice that $p(x)$ will be zero for other values of X. Then, Distribution function of X is given by

X	$F(x) = P[X \le x]$
0	$\frac{1}{4}$
1	$\frac{1}{4} + \frac{1}{2} = \frac{3}{4}$
2	$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

Here, for the last value, i.e. for X = 2, we have F(x) =1.

Also, if we take a value beyond 2 say 4, then we get

$$\begin{aligned} F(4) &= P[X \le 4] \\ &= P[X = 4] + P[X = 3] + P[X \le 2] \\ &= 0 + 0 + 1 = 1. \end{aligned}$$

Example 7: A random variable X has the following probability function:

X	0	1	2	3	4	5	6	7
p(x)	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{100}$	$\frac{1}{50}$	$\frac{17}{100}$

Determine the distribution function of X.

Solution: Here,

$F(0) = P[X \le 0] = P[X = 0] = 0$,

$F(1) = P[X \le 1] = P[X = 0] + P\,[X = 1] = 0 + \frac{1}{10} = \frac{1}{10}$,

$F(2) = P[X \le 2] = P[X = 0] + P\,[X = 1] + [X = 2] = 0 + \frac{1}{10} + \frac{1}{5} = \frac{3}{10}$,

and so on. Thus, the distribution function F(x) of X is given in the following table:

X	F(x) = P[X ≤ x]
0	0
1	$\frac{1}{10}$
2	$\frac{3}{10}$
3	$\frac{3}{10} + \frac{1}{5} = \frac{1}{2}$
4	$\frac{1}{2} + \frac{3}{10} = \frac{4}{5}$
5	$\frac{4}{5} + \frac{1}{100} = \frac{81}{100}$
6	$\frac{81}{100} + \frac{1}{50} = \frac{83}{100}$
7	$\frac{83}{100} + \frac{17}{100} = 1$

Continuous Distribution Function

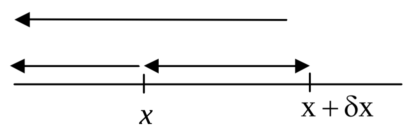
Distribution function of a continuous random variable is called the continuous distribution function or cumulative distribution function (c.d.f.).

Let X be a continuous random variable having the probability density function f(x), as defined in the last section of this unit, then the distribution function F(x) is given by

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx .$$

Also, in the last section, we have defined the p.d.f. f(x) as

$$f(x) = \lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x},$$



$$\therefore f(x) = \lim_{\delta x \to 0} \frac{P[X \le x + \delta x] - P[X \le x]}{\delta x}$$

$$\Rightarrow f(x) = \lim_{\delta x \to 0} \frac{F(x + \delta x) - F(x)}{\delta x},$$

$$\Rightarrow f(x) = \text{Derivative of } F(x) \text{ with respect to } x \quad \begin{bmatrix} \text{By definition of} \\ \text{the derivative} \end{bmatrix}$$

$$\Rightarrow f(x) = F'(x)$$

$$\Rightarrow f(x) = \frac{d}{dx}\left(F(x)\right)$$

$$\Rightarrow dF(x) = f(x)\,dx$$

Here, dF(x) is known as the probability differential.

So, $F(x) = \int_{-\infty}^{x} f_X(x)\,dx$ and $F'(x) = f(x)$.

Example 8: The diameter 'X' of a cable is assumed to be a continuous random variable with p.d.f. $f(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$.

Obtain the c.d.f. of X.

Solution: For $0 \le x \le 1$, the c.d.f. of X is given by

$$F(x) = P[X \le x] = \int_{0}^{x} f(x)\,dx = \int_{0}^{x} 6x(1-x)\,dx$$

$$= 6\int_{0}^{x} \left(x - x^2\right) dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^x = 3x^2 - 2x^3$$

$\therefore$ The c.d.f. of X is given by

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

Remark 4: In the above example, F(x) is taken as 0 for $x<0$ since p(x) = 0 for $x<0$; and F(x) is taken as 1 for $x>1$ since F(1) = 1 and therefore,

for $x>1$ also F(x) will remain 1.

Now, you can try the following exercises.

E 5) A random variable X has the following probability distribution:

X	0	1	2	3	4	5	6	7	8
p(*x*)	k	3k	5k	7k	9k	11k	13k	15k	17k

i) Determine the value of k.

ii) Find the distribution function of X.

E 6) Let X be continuous random variable with p.d.f. given by.

$$f(x)=\begin{cases}\frac{x}{2}, & 0\le x<1\\ \frac{1}{2}, & 1\le x<2\\ \frac{1}{2}(3-x), & 2\le x<3\\ 0, & \text{elsewhere.}\end{cases}.$$

Determine F(x), the c.d.f. of X.

5.6 SUMMARY

Following main points have been covered in this unit of the course:

1) A **random variable** is a function whose domain is a set of possible outcomes and range is a sub-set of the set of reals and has the following properties:

   i) Each particular value of the random variable can be assigned some probability.

   ii) Sum of all the probabilities associated with all the different values of the random variable is unity.

2) A random variable is said to be **discrete random variable** if it has either a finite number of values or a countable number of values, i.e. the values can be arranged in a sequence.

3) If a random variable is such that its values cannot be arranged in a sequence, it is called **continuous random variable**. So, a random variable is said to be continuous if it can take all the possible real (i.e. integer as well as fractional) values between two certain limits.

4) Let X be a discrete r.v. which take on the values x_1, x_2, ... and let $P\left[X = x_i\right] = p(x_i)$. The function $p(x_i)$ is called **probability mass function** of X satisfying $p(x_i) \geq 0$ and $\sum_i p\left(x_i\right) = 1$. The set $\left\{\left(x_1, p\left(x_1\right)\right), \left(x_2, p\left(x_2\right)\right), \ldots\right\}$ specifies the **probability distribution** of discrete r.v. X.

5) Let X be a continuous random variable and $f(x)$ be a continuous function of x. Suppose $(x, x+\delta x)$ be an interval of length δx. Then $f(x)$ defined by $\lim_{\delta x \to 0} \frac{P[x \leq X \leq x + \delta x]}{\delta x} = f\left(x\right)$ is called the **probability density function** of X.

   Probability density function has the same properties as that of probability mass function i.e. $f(x) \geq 0$ and $\int_R f\left(x\right) dx = 1$, where integral has been taken over the entire range R of values of X.

6) A function F defined for all values of a random variable X by $F(x) = P[X \leq x]$ is called the **distribution function**. It is also known as the **cumulative distribution function (c.d.f.)** of X. The domain of the distribution function is a set of real numbers and its range is [0, 1]. Distribution function of a discrete random variable X is said to be **discrete distribution function** and is given by $\left\{\left(x_1, F\left(x_1\right)\right), \left(x_2, F\left(x_2\right)\right), \ldots\right\}$. Distribution function of a continuous random variable X having the probability density function $f(x)$ is said to be **continuous distribution function** and is given by

   $F(x) = P[X \leq x] = \int_{-\infty}^{x} f\left(x\right) dx$.

   Derivative of $F(x)$ with respect to x is $f\left(x\right)$, i.e. $F'\left(x\right) = f\left(x\right)$.

5.7 SOLUTIONS/ANSWERS

E 1) Let X be the number of bad articles drawn.

$\therefore$ X can take the values 0, 1, 2 with

P[X = 0] = P[No bad article]

= P[Drawing 2 articles from 5 good articles and zero article from 2 bad articles]

$$= \frac{{}^5C_2 \times {}^2C_0}{{}^7C_2} = \frac{5\times4\times1}{7\times6} = \frac{10}{21},$$

P[X = 1] = P[One bad article and 1 good article]

$$= \frac{{}^2C_1 \times {}^5C_1}{{}^7C_2} = \frac{2\times5\times2}{7\times6} = \frac{10}{21}, \text{ and}$$

P[X = 2] = P[Two bad articles and no good article]

$$= \frac{{}^2C_2 \times {}^5C_0}{{}^7C_2} = \frac{1\times1\times2}{7\times6} = \frac{1}{21}$$

$\therefore$ Probability distribution of number of bad articles is:

X	0	1	2
p(x)	$\frac{10}{21}$	$\frac{10}{21}$	$\frac{1}{21}$

E2) As $Y = X^2 + 2X$,

$\therefore$ For X = 0, Y = 0 + 0 = 0;

For X = 1, $Y = 1^2 + 2(1) = 3$;

For X = 2, $Y = 2^2 + 2(2) = 8$; and

For X = 3, $Y = 3^2 + 2(3) = 15$.

Thus, the values of Y are 0, 3, 8, 15 corresponding to the values 0, 1, 2, 3 of X and hence

$P[Y = 0] = P[X = 0] = \frac{1}{10}$, $P[Y = 3] = P[X = 1] = \frac{3}{10}$,

$P[Y = 8] = P[X = 2] = \frac{1}{2}$ and $P[Y = 15] = P[X = 3] = \frac{1}{10}$.

$\therefore$ The probability distribution of Y is

Y	0	3	8	15
p(y)	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{10}$

E3) Let X be the number of red balls drawn.

$\therefore$ X can take the values 0, 1, 2, 3.

Let W_i be the event that i^{th} draw gives white ball and R_i be the event that i^{th} draw gives red ball.

$\therefore$ $P[X = 0] = P[\text{No Red ball}] = P[W_1 \cap W_2 \cap W_3]$

$= P(W_1).\ P(W_2).\ P(W_3)$

[$\because$ balls are drawn with replacement and hence the draws are independent]

$$= \frac{3}{7} \times \frac{3}{7} \times \frac{3}{7} = \frac{27}{343}$$

P[X = 1] = P[One red and two white]

$$= P\left[\left(R_1 \cap W_2 \cap W_3\right) \text{or} \left(W_1 \cap R_2 \cap W_3\right) \text{or} \left(W_1 \cap W_2 \cap R_3\right)\right]$$

$$= P\left[R_1 \cap W_2 \cap W_3\right] + P\left[W_1 \cap R_2 \cap W_3\right] + P\left[W_1 \cap W_2 \cap R_3\right]$$

$$= P[R_1]P\left[W_2\right]P\left[W_3\right] + P\left[W_1\right]P\left[R_2\right]P\left[W_3\right] + P\left[W_1\right]P\left[W_2\right]P\left[R_3\right]$$

$$= \frac{4}{7} \times \frac{3}{7} \times \frac{3}{7} + \frac{3}{7} \times \frac{4}{7} \times \frac{3}{7} + \frac{3}{7} \times \frac{3}{7} \times \frac{4}{7} = 3 \times \frac{4}{7} \times \frac{3}{7} \times \frac{3}{7} = \frac{108}{343},$$

P[X = 2] = P[Two red and one white]

$$= P\left[\left(R_1 \cap R_2 \cap W_3\right) \text{or} \left(R_1 \cap W_2 \cap R_3\right) \text{or} \left(W_1 \cap R_2 \cap R_3\right)\right]$$

$$= P[R_1]P\left[R_2\right]P\left[W_3\right] + P\left[R_1\right]P\left[W_2\right]P\left[R_3\right] + P\left[W_1\right]P\left[R_2\right]P\left[R_3\right]$$

$$= \frac{4}{7} \times \frac{4}{7} \times \frac{3}{7} + \frac{4}{7} \times \frac{3}{7} \times \frac{4}{7} + \frac{3}{7} \times \frac{4}{7} \times \frac{4}{7} = 3 \times \frac{4}{7} \times \frac{4}{7} \times \frac{3}{7} = \frac{144}{343}.$$

P[X = 3] = P [Three red balls]

$$= P[R_1 \cap R_2 \cap R_3] = P(R_1)\, P(R_2)\, P(R_3) = \frac{4}{7} \times \frac{4}{7} \times \frac{4}{7} = \frac{64}{343}.$$

$\therefore$ Probability distribution of the number of red balls is

X	0	1	2	3
$p(x)$	$\frac{27}{343}$	$\frac{108}{343}$	$\frac{144}{343}$	$\frac{64}{343}$

E4) As $f(x)$ is p.d.f.,

$$\therefore \int_{1500}^{2500} \frac{A}{x^3}\,dx = 1 \Rightarrow A \int_{1500}^{2500} x^{-3}\,dx = 1 \Rightarrow A\left[\frac{x^{-2}}{-2}\right]_{1500}^{2500} = 1$$

$$\Rightarrow -\frac{A}{2}\left[\frac{1}{(2500)^2} - \frac{1}{(1500)^2}\right] = 1 \Rightarrow -\frac{A}{20000}\left[\frac{1}{625} - \frac{1}{225}\right] = 1$$

$$\Rightarrow -\frac{A}{20000}\left[\frac{9-25}{5625}\right] = 1 \Rightarrow 16A = 5625 \times 20000$$

$$\Rightarrow \quad A = \frac{5625 \times 20000}{16} = 5625 \times 1250 = 7031250.$$

Now, $P[1600 \le X \le 2000] = \int_{1600}^{2000} f(x)\,dx = A \int_{1600}^{2000} \frac{1}{x^3}\,dx$

$$= -\frac{A}{2}\left[\frac{1}{x^2}\right]_{1600}^{2000} = -\frac{A}{2}\left[\frac{1}{(2000)^2} - \frac{1}{(1600)^2}\right]$$

$$= -\frac{A}{20000}\left[\frac{1}{400} - \frac{1}{256}\right] = -\frac{A}{20000}\left[\frac{16-25}{6400}\right]$$

$$= \frac{9 \times 7031250}{20000 \times 6400} = \frac{2025}{4096}$$

E5) i) As the given distribution is probability distribution,

$\therefore$ sum of all the probabilities = 1

$\Rightarrow$ k + 3k + 5k + 7k + 9k + 11k + 13k + 15k + 17k = 1

$\Rightarrow 81\text{ k} = 1 \Rightarrow \text{k} = \frac{1}{81}$

ii) The distribution function of X is given in the following table:

X	$F(x) = P[X \le x]$
0	$k = \frac{1}{81}$
1	$k + 3k = 4k = \frac{4}{81}$
2	$4k + 5k = 9k = \frac{9}{81}$
3	$9k + 7k = 16k = \frac{16}{81}$
4	$16k + 9k = 25k = \frac{25}{81}$
5	$25k + 11k = 36k = \frac{36}{81}$
6	$36k + 13k = 49k = \frac{49}{81}$
7	$49k + 15k = 64k = \frac{64}{81}$
8	$64k + 17k = 81k = 1$

E6) For $x < 0$,

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{x} 0\,dx = 0 \quad [\because f(x) = 0 \text{ for } x < 0].$$

For $0 \le x < 1$,

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{x} f(x)\,dx \quad [\because 0 \le x < 1]$$

$$= 0 + \int_{0}^{x} \frac{x}{2}\,dx \qquad \left[\because f(x) = \frac{x}{2} \text{ for } 0 \le x < 1\right]$$

$$= \frac{1}{2}\left[\frac{x^2}{2}\right]_0^x = \frac{x^2}{4}.$$

For $1 \le x < 2$,

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

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

$$= 0 + \int_{0}^{1} \frac{x}{2}\,dx + \int_{1}^{x} \frac{1}{2}\,dx = \frac{1}{4}\left[x^2\right]_0^1 + \frac{1}{2}[x]_1^x$$

$$= \frac{1}{4} + \frac{x}{2} - \frac{1}{2} = \frac{1}{4}(2x - 1)$$

For $2 \le x < 3$,

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

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

$$= 0 + \int_{0}^{1} \frac{x}{2}\,dx + \int_{1}^{2} \frac{1}{2}\,dx + \int_{2}^{x} \frac{1}{2}(3 - x)\,dx$$

$$= \left[\frac{x^2}{4}\right]_0^1 + \frac{1}{2}[x]_1^2 + \frac{1}{2}\left[3x - \frac{x^2}{2}\right]_2^x$$

$$= \frac{1}{4} + \frac{1}{2} + \frac{1}{2}\left[\left(3x - \frac{x^2}{2}\right) - (6 - 2)\right]$$

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

For $3 \le x < \infty$,

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

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{2} f(x)\,dx + \int_{2}^{3} f(x)\,dx \int_{3}^{x} f(x)\,dx$$

$$= 0 + \int_{0}^{1} \frac{x}{2}\,dx + \int_{1}^{2} \frac{1}{2}\,dx + \int_{2}^{3} \frac{1}{2}(3-x)\,dx + \int_{3}^{x} 0\,dx$$

$$= \left[\frac{x^2}{4}\right]_0^1 + \left[\frac{x}{2}\right]_1^2 + \frac{1}{2}\left[3x - \frac{x^2}{2}\right]_2^3 + 0$$

$$= \frac{1}{4} + \left(1 - \frac{1}{2}\right) + \frac{1}{2}\left[\left(9 - \frac{9}{2}\right) - (6-2)\right]$$

$$= \frac{1}{4} + \frac{1}{2} + \frac{1}{2}\left(\frac{9}{2} - 4\right)$$

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

Hence, the distribution function is given by:

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

UNIT 6 BIVARIATE DISCRETE RANDOM VARIABLES

Structure

6.1 INTRODUCTION

In Unit 5, you have studied one-dimensional random variables and their probability mass functions, density functions and distribution functions. There may also be situations where we have to study two-dimensional random variables in connection with a random experiment. For example, we may be interested in recording the number of boys and girls born in a hospital on a particular day. Here, 'the number of boys' and 'the number of girls' are random variables taking the values 0, 1, 2, … and both these random variables are discrete also.

In this unit, we concentrate on the two-dimensional discrete random variables defining them in Sec. 6.2. The joint, marginal and conditional probability mass functions of two-dimensional random variable are described in Sec. 6.3. The distribution function and the marginal distribution function are discussed in Sec. 6.4.

Objectives

A study of this unit would enable you to:

- define two-dimensional discrete random variable;
- specify the joint probability mass function of two discrete random variables;
- obtain the marginal and conditional distributions for two-dimensional discrete random variable;
- define two-dimensional distribution function;
- define the marginal distribution functions; and
- solve various practical problems on bivariate discrete random variables.

6.2 BIVARIATE DISCRETE RANDOM VARIABLES

In Unit 5, the concept of single-dimensional random variable has been studied in detail. Proceeding in analogy with the one-dimensional case, concept of two-dimensional discrete random variables is discussed in the present unit.

A situation where two-dimensional discrete random variable needs to be studied has already been given in Sec. 6.1 of this unit. To describe such situations mathematically, the study of two random variables is introduced.

Definition: Let X and Y be two discrete random variables defined on the sample space S of a random experiment then the function (X, Y) defined on the same sample space is called a two-dimensional discrete random variable. In others words, (X, Y) is a two-dimensional random variable if the possible values of (X, Y) are finite or countably infinite. Here, each value of X and Y is represented as a point (x, y) in the xy-plane.

As an illustration, let us consider the following example:

Let three balls b_1, b_2, b_3 be placed randomly in three cells. The possible outcomes of placing the three balls in three cells are shown in Table 6.1.

Table 6.1 : Possible Outcomes of Placing the Three Balls in Three Cells

Arrangement Number	Placement of the Balls in		
	Cell 1	Cell 2	Cell 3
1	b_1	b_2	b_3
2	b_1	b_3	b_2
3	b_2	b_1	b_3
4	b_2	b_3	b_1
5	b_3	b_1	b_2
6	b_3	b_2	b_1
7	b_1,b_2	b_3	-
8	b_1,b_2	-	b_3
9	-	b_1,b_2	b_3
10	b_1,b_3	b_2	-
11	b_1,b_3	-	b_2
12	-	b_1,b_3	b_2
13	b_2,b_3	b_1	-
14	b_2,b_3	-	b_1
15	-	b_2,b_3	b_1
16	b_1	b_2,b_3	-
17	b_1	-	b_2,b_3
18	-	b_1	b_2,b_3

19	b_2	b_3,b_1	-
20	b_2	-	b_3,b_1
21	-	b_2	b_3,b_1
22	b_3	b_1,b_2	-
23	b_3	-	b_1,b_2
24	-	b_3	b_1,b_2
25	b_1,b_2,b_3	-	-
26	-	b_1,b_2,b_3	-
27	-	-	b_1,b_2,b_3

Now, let X denote the number of balls in Cell 1 and Y be the number of cells occupied. Notice that X and Y are discrete random variables where X take on the values 0, 1, 2, 3 ($\because$ number of balls in Cell 1 may be 0 or 1 or 2 or 3) and Y take on the values 1, 2, 3 ($\because$ number of occupied cells may be 1 or 2 or 3). The possible values of two-dimensional random variable (X, Y), therefore, are all ordered pairs of the values x and y of X and Y, respectively, i.e. are (0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3).

Now, to each possible value $\left(x_i, y_j\right)$ of (X, Y), we can associate a number $p\left(x_i, y_j\right)$ representing $P\left(X = x_i, Y = y_j\right)$ as discussed in the following section of this unit.

6.3 JOINT, MARGINAL AND CONDITIONAL PROBABILITY MASS FUNCTIONS

Let us again consider the example discussed in Sec. 6.2. In this example, we have obtained all possible values of (X, Y), where X is the number of balls in Cell 1 and Y be the number of occupied cells. Now, let us associate numbers $p\left(x_i, y_j\right)$ representing $P\left[X = x_i, Y = y_j\right]$ as follows:

p(0, 1) = P[X = 0, Y = 1] = P[no ball in Cell 1 and 1 cell occupied] = P[Arrangement numbers 26, 27] = $\frac{2}{27}$
p(0, 2) = P[X = 0, Y = 2] = P[no ball in Cell 1 and 2 cells occupied] = P[Arrangement numbers 9, 12, 15, 18, 21, 24] = $\frac{6}{27}$
p(0, 3) = P[X = 0, Y = 3] = P[no ball in Cell 1 and 3 cells occupied] = P[Impossible event] = 0
p(1, 1) = P[X = 1, Y = 1] = P[one ball in Cell 1 and 1 cell occupied] = P[Impossible event] = 0

$p(1, 2) = P[X = 1, Y = 2]$ = P[one ball in Cell 1 and 2 cells occupied] = P[Arrangement numbers 16, 17, 19, 20, 22, 23] $= \frac{6}{27}$
$p(1, 3) = P[X = 1, Y = 3]$ = P[one ball in Cell 1 and 3 cells occupied] = P[Arrangement numbers1 to 6] = $\frac{6}{27}$
$p(2, 1) = P[X = 2, Y = 1]$ = P[two balls in Cell 1 and 1 cell occupied] = P[Impossible event] = 0
$p(2, 2) = P[X = 2, Y = 2]$ = P[two balls in Cell 1 and 2 cells occupied] = P[Arrangement numbers 7, 8, 10, 11, 13, 14] $= \frac{6}{27}$
$p(2, 3) = P[X = 2, Y = 3]$ = P[two balls in Cell 1 and 3 cells occupied] = P[Impossible event] = 0
$p(3, 1) = P[X = 3, Y = 1]$ = P[three balls in Cell 1 and 1 cell occupied] = P[Arrangement number 25] = $\frac{1}{27}$
$p(3, 2) = P[X = 3, Y = 2]$ = P[three balls in Cell 1 and 2 cells occupied] = P[Impossible event] = 0
$p(3, 3) = P[X = 3, Y = 3]$ = P[three balls in Cell 1 and 3 cells occupied] = P[Impossible event] = 0

The values of (X, Y) together with the number associated as above constitute what is known as joint probability distribution of (X, Y) which can be written in the tabular form also as shown below:

X \ Y	1	2	3	Total
0	2 / 27	6 / 27	0	8 / 27
1	0	6 / 27	6 / 27	12 / 27
2	0	6 / 27	0	6 / 27
3	1 / 27	0	0	1 / 27
Total	3 / 27	18 / 27	6 / 27	1

We are now in a position to define joint, marginal and conditional probability mass functions.

Joint Probability Mass Function

Let (X, Y) be a two-dimensional discrete random variable. With each possible outcome $\left(x_i, y_j\right)$, we associate a number $p\left(x_i, y_j\right)$ representing $P\left[X = x_i, Y = y_j\right]$ or $P\left[X = x_i \cap Y = y_j\right]$ and satisfying following conditions:

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

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

The function p defined for all $\left(x_i, y_j\right)$ is in analogy with one-dimensional case and called the **joint probability mass function of X and Y.** It is usually represented in the form of the table as shown in the example discussed above.

Marginal Probability Function

Let (X, Y) be a discrete two-dimensional random variable which take up finite or countably infinite values $\left(x_i, y_j\right)$. For each such two-dimensional random variable (X, Y), we may be interested in the probability distribution of X or the probability distribution of Y, individually.

Let us again consider the example of the random placement of three balls in there cells wherein X and Y are the discrete random variables representing "the number of balls in Cell 1" and "the number of occupied cells", respectively. Let us consider Table 6.1 showing the joint distribution of (X, Y). From this table, let us take up the row totals and column totals. The row totals in the table represent the probability distribution of X and the column totals represent the probability distribution of Y, individually. That is,

$$P[X = 0] = \frac{2}{27} + \frac{6}{27} + 0 = \frac{8}{27}$$

$$P[X = 1] = 0 + \frac{6}{27} + \frac{6}{27} = \frac{12}{27}$$

$$P[X = 2] = 0 + \frac{6}{27} + 0 = \frac{6}{27}$$

$$P[X = 3] = \frac{1}{27} + 0 + 0 = \frac{1}{27} \text{ and}$$

$$P[Y = 1] = \frac{2}{27} + 0 + 0 + \frac{1}{27} = \frac{3}{27}$$

$$P[Y = 2] = \frac{6}{27} + \frac{6}{27} + \frac{6}{27} + 0 = \frac{18}{27}$$

$$P[Y = 3] = 0 + \frac{6}{27} + 0 + 0 = \frac{6}{27}$$

These distributions of X and Y, individually, are called the **marginal probability distributions** of X and Y, respectively.

So, if (X, Y) is a discrete two-dimensional random variable which take up the values (x_i, y_j), then the probability distribution of X is determined as follows:

$$p(x_i) = P[X = x_i]$$

$$= P[(X = x_i \cap Y = y_1) \text{ or } (X = x_i \cap Y = y_2) \text{ or} \ldots]$$

$$= P[X = x_i \cap Y = y_1] + P[X = x_i \cap Y = y_2] + P[X = x_i \cap Y = y_3] + \ldots$$

$$= \sum_j P[X = x_i \cap Y = y_j]$$

$$= \sum_j p(x_i, y_j) \left[\begin{array}{l} \because\ p(x_i, y_j), \text{the joint probability mass} \\ \text{function, is } P[X = x_i \cap Y = y_j] \end{array} \right]$$

which is known as the marginal probability mass function of X. Similarly, the probability distribution of Y is

$$p(y_j) = P[Y = y_j]$$

$$= P[X = x_1 \cap Y = y_j] + P[X = x_2 \cap Y = y_j] + \ldots$$

$$= \sum_i P[X = x_i \cap Y = y_j]$$

$$= \sum_i p(x_i, y_j)$$

and is known as the marginal probability mass function of Y.

Conditional Probability Mass Function

Let (X, Y) be a discrete two-dimensional random variable. Then the conditional probability mass function of X, given $Y = y$ is defined as

$$p(x \mid y) = P[X = x \mid Y = y]$$

$$= \frac{P[X = x \cap Y = y]}{P[Y = y]}, \text{ provided } P[Y = y] \neq 0$$

$$\left[\because P[A \mid B] = \frac{P[A \cap B]}{P[B]}, P(B) \neq 0 \right]$$

Similarly, the conditional probability mass function of Y, given X = x, is defined as

$$p(y \mid x) = P[Y = y \mid X = x] = \frac{P[Y = y \cap X = x]}{P[X = x]}$$

Let us again consider the example as already discussed in this section. Suppose, we are interested in finding the conditional probability mass function of X given Y = 2. Then the conditional probabilities are found separately for each value of X given Y = 2. That is, we proceed as follows:

$$P[X=0\,|\,Y=2]=\frac{P[X=0\cap Y=2]}{P[Y=2]}=\frac{\frac{6}{27}}{\frac{18}{27}}=\frac{1}{3}$$

$$P[X=1\,|\,Y=2]=\frac{P[X=1\cap Y=2]}{P[Y=2]}=\frac{\frac{6}{27}}{\frac{18}{27}}=\frac{1}{3}$$

$$P[X=2\,|\,Y=2]=\frac{P[X=2\cap Y=2]}{P[Y=2]}=\frac{\frac{6}{27}}{\frac{18}{27}}=\frac{1}{3}$$

$$P[X=3\,|\,Y=2]=\frac{P[X=3\cap Y=2]}{P[Y=2]}=\frac{0}{\frac{18}{27}}=0$$

[Note that values of numerator and denominator in the above expressions have already been obtained while discussing the joint and marginal probability mass functions in this section of the unit.]

Independence of Random Variables

Two discrete random variables X and Y are said to be independent if and only if

$$P[X=x_i\cap Y=y_j]=P[X=x_i]P[Y=y_j]$$

[$\because$ two events A and B are independent if and only if $P(A\cap B) = P(A)\,P(B)$]

Example 1: The following table represents the joint probability distribution of the discrete random variable (X, Y):

X \ Y	1	2
1	0.1	0.2
2	0.1	0.3
3	0.2	0.1

Find **:**

i) The marginal distributions.

ii) The conditional distribution of X given Y = 1.

iii) P[(X + Y) < 4].

Solution:

i) To find the marginal distributions, we have to find the marginal totals, i.e. row totals and column totals as shown in the following table:

Y / X	1	2	$p(x)$ (Totals)
1	0.1	0.2	0.3
2	0.1	0.3	0.4
3	0.2	0.1	0.3
$p(y)$ (Totals)	0.4	0.6	1

Thus, the marginal probability distribution of X is

X	1	2	3
p(x)	0.3	0.4	0.3

and the marginal probability distribution of Y is

Y	1	2
P(y)	0.4	0.6

ii) As $P[X=1|Y=1]=\frac{P[X=1,Y=1]}{P[Y=1]}=\frac{0.1}{0.4}=\frac{1}{4}$,

$$P[X=2|Y=1]=\frac{P[X=2,Y=1]}{P[Y=1]}=\frac{0.1}{0.4}=\frac{1}{4} \text{ and}$$

$$P[X=3|Y=1]=\frac{P[X=3\cap Y=1]}{P[Y=1]}=\frac{0.2}{0.4}=\frac{1}{2},$$

$\therefore$ The conditional distribution of X given Y = 1 is

X	1	2	3
$P[X=x\,|\,Y=1]$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

(iii) As the values of (X, Y) which satisfy X + Y < 4 are (1, 1), (1, 2) and (2, 1) only.

$$\therefore P\left[(X+Y)<4\right]=P[X=1,Y=1]+P[X=1,Y=2]+P[X=2,Y=1]$$

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

Example 2: Two discrete random variables X and Y have

$P[X = 0,\ Y = 0] = \frac{2}{9}$, $P[X = 0,\ Y = 1] = \frac{1}{9}$, $P[X=1,Y=0]=\frac{1}{9}$, and

$P[X=1, Y=1]=\frac{5}{9}$. Examine whether X and Y are independent?

Solution: Writing the given distribution in tabular form as follows:

X \ Y	0	1	$p(x)$
0	2/9	1/9	3/9
1	1/9	5/9	6/9
$p(y)$	3/9	6/9	1

$\therefore\ P[X=0]=\frac{3}{9},\ P[X=1]=\frac{6}{9},$

$P[Y=0]=\frac{3}{9},\ P[Y=1]=\frac{6}{9}$

Now $P[X=0]P[Y=0]=\frac{3}{9}\times\frac{3}{9}=\frac{1}{9}$

But $P[X=0, Y=0]=\frac{2}{9}$

$\therefore\ P[X=0, Y=0]\neq P[X=0]P[Y=0]$

Hence X and Y are not independent

[**Note**: If $P[X=x, Y=y]=P[X=x]P[Y=y]$ for each possible value of X and Y, only then X and Y are independent.]

Here are two exercises for you.

E1) The joint probability distribution of a pair of random variables is given by the following table:

X \ Y	1	2	3
1	1/12	0	1/18
2	1/6	1/9	1/4
3	0	1/5	2/15

i) Evaluate marginal distribution of X.

ii) Evaluate conditional distribution of Y given X = 2

iii) Obtain $P[X+Y<5]$.

E2) For the following joint probability distribution of (X, Y),

X \ Y	1	2	3
1	$1/20$	$1/10$	$1/10$
2	$1/20$	$1/10$	$1/10$
3	$1/10$	$1/10$	$1/20$
4	$1/10$	$1/10$	$1/20$

i) find the probability that Y = 2 given that X = 4,

ii) find the probability that Y = 2, and

iii) examine if the two events X = 4 and Y = 2 are independent.

6.4 JOINT AND MARGINAL DISTRIBUTION FUNCTIONS FOR DISCRETE RANDOM VARIABLES

Two-Dimensional Joint Distribution Function

In analogy with the distribution function $F(x) = P[X \le x]$ of one-dimensional random variable X discussed in Unit 5 of this course, the distribution function of the two-dimensional random variable (X, Y) for all real x and y is defined as

$$F(x, y) = P[X \le x, Y \le y]$$

Marginal Distribution Functions

Let (X, Y) be a two-dimensional discrete random variable having $F(x, y)$ as its distribution function. Now the **marginal distribution function of X** is defined as

$$F(x) = P[X \le x]$$

$$= P[X \le x, Y = y_1] + P[X \le x, Y = y_2] + \ldots$$

$$= \sum_j P[X \le x, Y = y_j]$$

Similarly, the **marginal distribution function of Y** is defined as

$$F(y) = P[Y \le y]$$

$$= P[X = x_1, Y \le y] + P[X = x_2, Y \le y] + \ldots$$

$$= \sum_i P[X = x_i, Y \le y]$$

Example 3: Considering the probability distribution function given in Example 1, find

i) F(2, 2), F(3,2)

ii) $F_X(3)$

iii) $F_Y(1)$

Solution:

i) $F(2,2) = P[X \le 2, Y \le 2]$

$= P[X = 2, Y \le 2] + P[X = 1, Y \le 2]$

$= P[X = 2, Y = 2] + P[X = 2, Y = 1] + P[X = 1, Y = 2]$

$+ P[X = 1, Y = 1]$

= 0.3 + 0.1 + 0.2 + 0.1 = 0.7

$F(3,2) = P[X \le 3, Y \le 2]$

$= P[X \le 2, Y \le 2] + P[X = 3, Y \le 2]$

$= 0.7 + P[X = 3, Y \le 2]$ $\left[\begin{array}{l}\because \text{first term on R.H.S. has been} \\ \text{obtained in part (i) of this example}\end{array}\right]$

$= 0.7 + P[X = 3, Y = 2] + P[X = 3, Y = 1] = 0.7 + 0.1 + 0.2 = 1$

ii) $F_X(3) = P[X \le 3]$

$= P[X \le 3, Y = 1] + P[X \le 3, Y = 2]$

$= P[X = 3, Y = 1] + P[X = 2, Y = 1] + P[X = 1, Y = 1]$

$+ P[X = 3, Y = 2] + P[X = 2, Y = 2] + P[X = 1, Y = 2]$

= 0.2 + 0.1 + 0.1 +0.1+ 0.3 + 0.2 = 1

iii) $F_Y(1) = P[Y \le 1]$

$= P[X = 1, Y \le 1] + P[X = 2, Y \le 1] + P[X = 3, Y \le 1]$

$= P[X = 1, Y = 1] + P[X = 2, Y = 1] + P[X = 3, Y = 1]$

= 0.1 + 0.1 + 0.2 = 0.4

Example 4: Find the joint and marginal distribution functions for the joint probability distribution given in Example 2.

Solution: For the joint distribution function, we have to find $F(x, y) = P[X \le x, Y \le y]$ for each x and y, i.e. we are to find F(0,0), F(0,1), F(1,0), F(1,1).

$$F(0,0) = P[X \le 0, Y \le 0] = P[X = 0, Y = 0] = \frac{2}{9}$$

$$F(0,1) = P[X \le 0, Y \le 1] = P[X = 0, Y = 0] + P[X = 0, Y = 1]$$

$$= \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$$

$$F(1,0) = P[X \le 1, Y \le 0] = P[X = 1, Y = 0] + P[X = 0, Y = 0]$$

$$= \frac{1}{9} + \frac{2}{9} = \frac{3}{9}$$

$$F(1,1) = P[X \le 1, Y \le 1] = P[X = 1, Y = 1] + P[X = 1, Y = 0]$$

$$+ P[X = 0, Y = 1] + P[X = 0, Y = 0]$$

$$= \frac{5}{9} + \frac{1}{9} + \frac{1}{9} + \frac{2}{9} = 1$$

Above distribution function $F(x, y)$ can be shown in the tabular form as follows:

	$Y \le 0$	$Y \le 1$
$X \le 0$	$2/9$	$3/9$
$X \le 1$	$3/9$	1

Marginal distribution function of X is obtained on finding $F(x) = P[X \le x]$ for each x, i.e. we have to obtain $F_X(0)$, $F_X(1)$.

$$F_X(0) = P[X \le 0] = P[X = 0]$$

$$= P[X = 0, Y = 0] + P[X = 0, Y = 1]$$

$$= \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$$

$$F_X(1) = P[X \le 1] = P[X \le 1, Y = 0] + P[X \le 1, Y = 1]$$

$$= P[X = 1, Y = 0] + P[X = 0, Y = 0]$$

$$+ P[X = 1, Y = 1] + P[X = 0, Y = 1]$$

$$= \frac{1}{9} + \frac{2}{9} + \frac{5}{9} + \frac{1}{9} = 1$$

$\therefore$ marginal distribution function of X is given as

X	F(x)
≤ 0	$3/9$
≤ 1	1

Similarly, marginal distribution function of Y can be obtained. [Do it yourself]

Here is an exercise for you.

E3) Obtain the joint and marginal distribution functions for the joint probability distribution given in **E 1).**

Now before ending this unit, let us summarizes what we have covered in it.

6.5 SUMMARY

In this unit we have covered the following main points:

1) If X and Y be two discrete random variables defined on the sample space S of a random experiment then the function (X, Y) defined on the same sample space is called a **two-dimensional discrete random variable**. In others words, (X, Y) is a two-dimensional random variable if the possible values of (X, Y) are finite or countably infinite.

2) A number $p(x_i, y_j)$ associated with each possible outcome (x_i, y_j) of a two-dimensional discrete random variable (X, Y) is called the **joint probability mass function of X and Y** if it satisfies the following conditions:

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

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

3) If (X,Y) is a discrete two-dimensional random variable which take up the values (x_i, y_j), then the probability distribution of X given by $p(x_i) = \sum_j p(x_i, y_j)$ is known as the **marginal probability mass function of X** and the probability distribution of Y given by $p(y_j) = \sum_i p(x_i, y_j)$ is known as the **marginal probability mass function of Y**.

4) The **conditional probability mass function of X given** $Y = y$ in case of a two-dimensional discrete random variable (X, Y) is defined as

$$p(x \mid y) = P[X = x \mid Y = y]$$

$$= \frac{P[X = x \cap Y = y]}{P[Y = y]}; \text{ and}$$

the **conditional probability mass function of Y, given X = x** is defined as

$$p(y \mid x) = P[Y = y \mid X = x]$$

$$= \frac{P[Y = y \cap X = x]}{P[X = x]}$$

5) Two discrete random variables X and Y are said to be independent if and only if

$$P[X = x_i \cap Y = y_j] = P[X = x_i]P[Y = y_j]$$

6.6 SOLUTIONS/ANSWERS

E1) Let us compute the marginal totals. Thus, the complete table with marginal totals is given as

X \ Y	1	2	3	p(x)
1	$\frac{1}{12}$	0	$\frac{1}{18}$	$\frac{1}{12}+0+\frac{1}{18}=\frac{5}{36}$
2	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{4}$	$\frac{1}{6}+\frac{1}{9}+\frac{1}{4}=\frac{19}{36}$
3	0	$\frac{1}{5}$	$\frac{2}{15}$	$0+\frac{1}{5}+\frac{2}{15}=\frac{1}{3}$
p(y)	$\frac{1}{4}$	$\frac{14}{45}$	$\frac{79}{180}$	1

Therefore,

i) Marginal distribution of X is

X	p(x)
1	5/36
2	19/36
3	1/3

ii) $$P[Y=1\,|\,X=2]=\frac{P[Y=1,X=2]}{P[X=2]}=\frac{1}{6}\times\frac{36}{19}=\frac{6}{19}$$

$$P[Y=2\,|\,X=2]=\frac{P[Y=2,X=2]}{P[X=2]}=\frac{1}{9}\times\frac{36}{19}=\frac{4}{19}$$

$$P[Y=3\,|\,X=2]=\frac{P[Y=3,X=2]}{P[X=2]}=\frac{1}{4}\times\frac{36}{19}=\frac{9}{19}$$

$\therefore$ The conditional distribution of Y given X = 2 is

Y	$P[Y=y\,\|\,X=2]$
1	6/19
2	4/19
3	9/19

iii) $P[X+Y<5]$

$$=P[X=1,Y=1]+P[X=1,Y=2]+P[X=1,Y=3]$$

$$+P[X=2,Y=1]+P[X=2,Y=2]+P[X=3,Y=1]$$

$$=\frac{1}{12}+0+\frac{1}{18}+\frac{1}{6}+\frac{1}{9}+0=\frac{15}{36}.$$

E2) First compute the marginal totals, then you will be able to find

i) $P[X=4]=\frac{1}{4}$, and hence

$$P[Y=2 \mid X=4]=\frac{P[Y=2,\ X=4]}{P[X=4]}=\frac{2}{5}$$

ii) $P[Y=2]=\frac{2}{5}$

iii) $P[X=4,Y=2]=\frac{1}{10},\ P[X=4]=\frac{1}{4}, P[Y=2]=\frac{2}{5}$

$$P[X = 4]\ P[Y = 2] = \frac{1}{4}\times\frac{2}{5} = \frac{1}{10}$$

$\therefore$ X= 4 and Y= 2 are independent

E3) To obtain joint distribution function $F(x,y)=P[X\le x, Y\le y]$, we have to obtain

$F(x,y)$ for each value of X and Y, i.e. we have to obtain

$F(1,1), F(1,2), F(1,3), F(2,1), F(2,2), F(2,3), F(3,1), F(3,2), F(3,3)$.

Then, the distribution function in tabular form is

	$Y\le 1$	$Y\le 2$	$Y\le 3$
$X\le 1$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{5}{36}$
$X\le 2$	$\frac{1}{4}$	$\frac{13}{36}$	$\frac{2}{3}$
$X\le 3$	$\frac{1}{4}$	$\frac{101}{180}$	1

Marginal distribution function of X is given as

X	F(x)
≤ 1	5/36
≤ 2	2/3
≤ 3	1

Marginal distribution function of Y is

Y	F(y)
≤ 1	1/4
≤ 2	101/180
≤ 3	1

UNIT 7 BIVARIATE CONTINUOUS RANDOM VARIABLES

Structure

7.1 INTRODUCTION

In Unit 6, we have defined the bivariate discrete random variable (X, Y), where X and Y both are discrete random variables. It may also happen that one of the random variables is discrete and the other is continuous. However, in most applications we deal only with the cases where either both random variables are discrete or both are continuous. The cases where both random variables are discrete have already been discussed in Unit 6. Here, in this unit, we are going to discuss the cases where both random variables are continuous.

In Unit 6, you have studied the joint, marginal and conditional probability functions and distribution functions in context of bivariate discrete random variables. Similar functions, but in context of bivariate continuous random variables, are discussed in this unit.

Bivariate continuous random variable is defined in Sec. 7.2. Joint and marginal density functions are described in Sec. 7.3. Sec. 7.4 deals with the conditional distribution and density functions. Independence of two continuous random variables is dealt with in Sec. 7.5. Some practical problems on two-dimensional continuous random variables are taken up in Sec. 7.6.

Objectives

A study of this unit would enable you to:

- define two-dimensional continuous random variable;
- specify the joint and marginal probability density functions of two continuous random variables;
- obtain the conditional density and distribution functions for two-dimensional continuous random variable;
- check the independence of two continuous random variables; and
- solve various practical problems on bivariate continuous random variables.

7.2 BIVARIATE CONTINUOUS RANDOM VARIABLES

Definition: If X and Y are continuous random variables defined on the sample space S of a random experiment, then (X, Y) defined on the same sample space S is called bivariate continuous random variable if (X, Y) assigns a point in xy-plane defined on the sample space S. Notice that it (unlike discrete random variable) assumes values in some non-countable set. Some examples of bivariate continuous random variable are:

1. A gun is aimed at a certain point (say origin of the coordinate system). Because of the random factors, suppose the actual hit point is any point (X, Y) in a circle of radius unity about the origin.

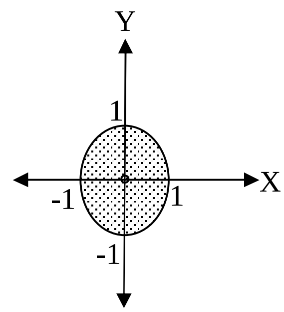


Fig.7.1: Actual Hit Point when a Gun is Aimed at a Certain Point

Then (X, Y) assumes all the values in the circle $\{(x, y) : x^2 + y^2 \le 1\}$ i.e. (X, Y) assumes all values corresponding to each and every point in the circular region as shown in Fig.7.1. Here, (X, Y) is bivariate continuous random variable.

2. (X, Y) assuming all values in the rectangle $\{(x, y) : a \le x \le b, c \le y \le d\}$ is a bivariate continuous random variable.

   Here, (X, Y) assumes all values corresponding to each and every point in the rectangular region as shown in Fig.7.2.

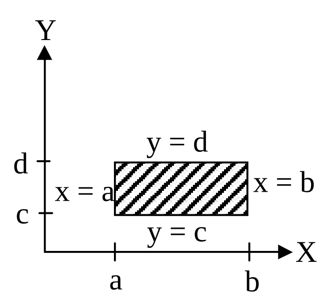


Fig.7.2: (X, Y) Assuming All Values in the Rectangle $\{(x,y) : a \le x \le b, c \le y \le d\}$

3. In a statistical survey, let X denotes the daily number of hours a child watches television and Y denotes the number of hours he/she spends on the studies. Here, (X, Y) is a two-dimensional continuous random variable.

7.3 JOINT AND MARGINAL DISTRIBUTION AND DENSITY FUNCTIONS

Two-Dimensional Continuous Distribution Function

The distribution function of a two-dimensional continuous random variable (X, Y) is a real-valued function and is defined as

$F(x, y) = P[X \le x, Y \le y]$ for all real x and y.

Notice that the above function is in analogy with one-dimensional continuous random variable case as studied in Unit 5 of the course.

Remark 1: $F(x, y)$ can also be written as $F_{X,Y}(x, y)$.

Joint Probability Density Function

Let (X, Y) be a continuous random variable assuming all values in some region R of the xy-plane. Then, a function $f(x, y)$ such that

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

is defined to be a joint probability density function.

As in the one-dimensional case, a joint probability density function has the following properties.

i) $f(x, y) \ge 0$

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

Remark 2:

As in the one-dimensional case, $f(x, y)$ does not represent the probability of anything. However, for positive δx and δy sufficiently small, $f(x, y)\delta x \delta y$ is approximately equal to

$P[x \le X \le x + \delta x, y \le Y \le y + \delta y]$.

In the one-dimensional case, you have studied that for positive δx sufficiently small $f(x)\delta x$ is approximately equal to $P[x \le X \le x + \delta x]$. So, the two-dimensional case is in analogy with the one-dimensional case.

Remark 3:

In analogy with the one-dimensional case [See Sec. 5.4 of Unit 5 of this course],

$f(x, y)$ can be written as $\lim\limits_{\substack{\delta x \to 0 \\ \delta y \to 0}} \frac{P[x \le X \le x + \delta x, y \le Y \le y + \delta y]}{\delta x \delta y}$

and is equal to

$\frac{\partial^2}{\partial x \partial y}(F(x, y))$, i.e. second order partial derivative with respect to x and y.

[See Sec. 5.5 of Unit 5 where $f(x) = \frac{d}{dx}(F(x))$]

Note: $\frac{\partial^2}{\partial x \partial y}(F(x,y))$ means first differentiate $F(x,y)$ partially w.r.t. y and then the resulting function w.r.t. x. When we differentiate a function partially w.r.t. one variable, then the other variable is treated as constant

For example, Let $F(x,y) = xy^3 + x^2y$

If we differentiate it partially w.r.t. y, we have

$\frac{\partial}{\partial y}(F(x,y)) = x(3y^2) + x^2.1$ [$\because$ here, x is treated as constant.]

If we now differentiate this resulting expression w.r.t. x, we have

$\frac{\partial^2}{\partial x \partial y}(F(x,y)) = 3y^2 + 2x$ [$\because$ here, y is treated as constant.]

Marginal Continuous Distribution Function

Let (X, Y) be a two-dimensional continuous random variable having $f(x,y)$ as its joint probability density function. Now, the marginal distribution function of the continuous random variable X is defined as

$$F(x) = P[X \le x]$$

$$= P[X \le x, y < \infty] \quad [\because \text{ for } X \le x, \text{ Y can take any real value}]$$

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

and the marginal distribution function of the continuous random variable Y is defined as

$$F(y) = P[Y \le y]$$

$$= P[Y \le y, X < \infty] \quad [\because \text{ for } Y \le y, \text{ X can take any real value}]$$

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

Marginal Probability Density Functions

Let (X, Y) be a two-dimensional continuous random variable having $F(x,y)$ and $f(x,y)$ as its distribution function and joint probability density function, respectively. Let $F(x)$ and $F(y)$ be the marginal distribution functions of X and Y, respectively. Then, the marginal probability density function of X is given as

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

or, it may also be obtained as $\frac{d}{dx}(F(x))$,

and the marginal probability density function of Y is given as

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

or

$$= \frac{d}{dy}(F(y))$$

7.4 CONDITIONAL DISTRIBUTION AND DENSITY FUNCTIONS

Conditional Probability Density Function

Let (X, Y) be a two-dimensional continuous random variable having the joint probability density function f(x, y). The conditional probability density function of Y given X = x is defined as

$$f(y \mid x) = \frac{f(x, y)}{f(x)}$$, where f(x) > 0 is the marginal density of X.

Similarly, the conditional probability density function of X given Y = y is defined to be

$$f(x \mid y) = \frac{f(x, y)}{f(y)}$$, where f(y) > 0 is the marginal density of Y.

As $f(y \mid x)$ and $f(x \mid y)$, though conditional yet, are the probability density functions, hence possess the properties of a probability density function. Properties of $f(y \mid x)$ are:

i) $f(y \mid x)$ is clearly ≥ 0

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

$$= \frac{1}{f(x)}\left[\int_{-\infty}^{\infty} f(x, y)\,dy\right]$$

$$= \frac{1}{f(x)}[f(x)] \qquad \left[\because \int_{-\infty}^{\infty} f(x, y)\,dy \text{ is the marginal probability density function of } X\right]$$

$$= 1$$

Similarly, $f(x \mid y)$ satisfies

i) $f(x \mid y) \geq 0$ and

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

Conditional Continuous Distribution Function

For a two-dimensional continuous random variable (X, Y), the conditional distribution function of Y given X = x is defined as

$$F(y \mid x) = P[Y \leq y \mid X = x]$$

$$= \int_{-\infty}^{y} f(y \mid x)\,dy \text{ , for all x such that } f(x) > 0;$$

and the conditional distribution function of X given Y = y is defined as

$$F(x \mid y) = P[X \leq x \mid Y = y]$$

$$= \int_{-\infty}^{x} f(x \mid y)\,dx \text{ , for all y such that f(y) > 0.}$$

7.5 STOCHASTIC INDEPENDENCE OF TWO CONTINUOUS RANDOM VARIABLES

You have already studied in Unit 3 of this course that independence of events is closely related to conditional probability, i.e. if events A and B are independent, then $P[A \mid B] = P[A]$, i.e. conditional probability of A is equal to the unconditional probability of A. Likewise independence of random variables is closely related to conditional distributions of random variables, i.e. two random variables X and Y with joint probability function $f(x, y)$ and marginal probability functions $f(x)$ and $f(y)$ respectively are said to be stochastically independent if and only if

i) $f(y \mid x) = f(y)$

ii) $f(x \mid y) = f(x)$.

Now, as defined in Sec. 7.4, we have

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

$\Rightarrow f(x, y) = f(x) f(y \mid x)$ [On cross-multiplying]

So, if X and Y are independent, then

$f(x, y) = f(x) f(y)$ [$\because \ f(y \mid x) = f(y)$]

Remark 4: The random variables, if independent, are actually stochastically independent but often the word "stochastically" is omitted.

Definition: Two random variables are said to be (stochastically) independent if and only if their joint probability density function is the product of their marginal density functions.

Let us now take up some problems on the topics covered so far in this unit.

7.6 PROBLEMS ON TWO-DIMENSIONAL CONTINUOUS RANDOM VARIABLES

Example 1: Let X and Y be two random variables. Then for

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

to be a joint density function, what must be the value of k ?

Solution: As $f(x,y)$ is the joint probability density function,

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

$$\Rightarrow \int_0^1\int_0^2 f(x,y)\,dy\,dx = 1 \qquad [\because 0<x<1,\ 0<y<2\ ]$$

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

$$\Rightarrow k\int_0^1\left[\int_0^2 (2x+y)\,dy\right]dx = 1$$

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

[Firstly the integral has been done w.r.t. y treating x as constant.]

$$\Rightarrow k\int_0^1\left[2x(2)+\frac{(2)^2}{2}-0\right]dx = 1$$

$$\Rightarrow k\int_0^1 (4x+2)\,dx = 1$$

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

$$\Rightarrow k\left[\frac{4}{2}+2-0\right] = 1 \Rightarrow 4k = 1 \Rightarrow k = \frac{1}{4}$$

Example 2: Let the joint density function of a two-dimensional random variable (X, Y) be:

$$f(x,y)=\begin{cases} x+y & \text{for } 0\le x<1 \text{ and } 0\le y<1 \\ 0, & \text{otherwise} \end{cases}$$

Find the conditional density function of Y given X.

Solution: The conditional density function of Y given X is $f(y\,|\,x)=\dfrac{f(x,y)}{f(x)}$,

where $f(x,y)$ is the joint density function, which is given; and $f(x)$ is the marginal density function which, by definition, is given by

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

$$=\int_0^1 f(x,y)\,dy \qquad [\because\ 0\le y<1]$$

$$=\int_0^1 (x+y)\,dy$$

$$=\left[xy+\frac{y^2}{2}\right]_0^1.$$

$$=\left[x(1)+\frac{(1)^2}{2}-0\right]=x+\frac{1}{2},\ 0\le x<1.$$

$\therefore$ the conditional density function of Y given X is

$$f(y\,|\,x)=\frac{f(x,y)}{f(x)}=\frac{x+y}{x+\frac{1}{2}},\ \text{for } 0\le x<1 \text{ and } 0\le y<1.$$

Example 3: Two-dimensional random variable (X, Y) have the joint density

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

i) Find $P[X<\frac{1}{2}\cap Y<\frac{1}{4}]$.

ii) Find the marginal and conditional distributions.

iii) Are X and Y independent?

Solution:

i) $P\left[X<\frac{1}{2}\cap Y<\frac{1}{4}\right]=\int_0^{\frac{1}{2}}\int_0^{\frac{1}{4}} f(x,y)\,dy\,dx=\int_0^{\frac{1}{2}}\int_0^{\frac{1}{4}} 8xy\,dy\,dx=\int_0^{\frac{1}{2}} 8x\left[\frac{y^2}{2}\right]_0^{\frac{1}{4}}dx$

$$=\int_0^{\frac{1}{2}} 8x\left[\frac{1}{16(2)}\right]dx \ = \int_0^{\frac{1}{2}} \frac{x}{4}dx \ = \frac{1}{4}\left[\frac{x^2}{2}\right]_0^{\frac{1}{2}}$$

$$=\frac{1}{4}\left[\frac{1}{8}\right]=\frac{1}{32}.$$

ii) Marginal density function of X is

$$f(x)=\int_x^1 f(x,y)\,dy \qquad [\because 0<x<y<1]$$

$$=\int_x^1 8xy\,dy \ = 8x\left[\frac{y^2}{2}\right]_x^1$$

$$=8x\left[\frac{1}{2}-\frac{x^2}{2}\right]=4x\left(1-x^2\right) \text{ for } 0<x<1$$

Marginal density function of Y is

$$f(y)=\int_0^y f(x,y)\,dx \qquad [\because 0<x<y]$$

$$=\int_0^y 8xy\,dx$$

$$=8y\left[\frac{x^2}{2}\right]_0^y=\frac{8y^3}{2}=4y^3 \quad \text{for} \quad 0<y<1$$

Conditional density function of X given Y(0 < Y < 1) is

$$f(x\,|\,y)=\frac{f(x,y)}{f(y)}$$

$$=\frac{8xy}{4y^3}=\frac{2x}{y^2},\ 0<x<y$$

Conditional density function of Y given X(0 < X < 1) is

$$f(y\,|\,x)=\frac{f(x,y)}{f(x)}$$

$$=\frac{8xy}{4x\left(1-x^2\right)} \ =\frac{2y}{\left(1-x^2\right)},\ x<y<1$$

iii) $f(x,y)=8xy,$

But $f(x)f(y)=4x\left(1-x^2\right)4y^3$

$$=16x\left(1-x^2\right)y^3$$

$$\therefore f(x, y) \neq f(x) f(y)$$

Hence, X and Y are not independent random variables.

Now, you can try some exercises.

E1) Let X and Y be two random variables. Then for

$$f(x, y) = \begin{cases} kxy & \text{for } 0 < x < 4 \text{ and } 1 < y < 5 \\ 0, & \text{otherwise} \end{cases}$$

to be a joint density function, what must be the value of k?

E2) If the joint p.d.f. of a two-dimensional random variable (X, Y) is given by

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

Then,

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

ii) Find the conditional density functions.

iii) Check for independence of X and Y.

E3) If (X, Y) be two-dimensional random variable having joint density function.

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

Find (i) $P[X < 1, Y < 3]$ (ii) $P[X < 1 | Y < 3]$

Now before ending this unit, let's summarize what we have covered in it.

7.7 SUMMARY

In this unit, we have covered the following main points:

1) If X and Y are continuous random variables defined on the sample space S of a random experiment, then (X, Y) defined on the same sample space S is called **bivariate continuous random variable** if (X, Y) assigns a point in xy-plane defined on the sample space S.

2) The distribution function of a two-dimensional continuous random variable (X, Y) is a real-valued function and is defined as

$$F(x, y) = P[X \leq x, Y \leq y] \text{ for all real x and y.}$$

3) A function $f(x, y)$ is called **joint probability density function** if it is such that

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

and satisfies

i) $f(x, y) \geq 0$

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

4) The **marginal distribution function** of the continuous random variable X is defined as

$$F(x) = P[X \leq x] = \int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(x, y)\,dy\right]dx,$$

and that of continuous random variable Y is defined as

$$F(y) = P[Y \leq y] = \int_{-\infty}^{y}\left[\int_{-\infty}^{\infty} f(x, y)\,dx\right]dy.$$

5) The **marginal probability density function** of X is given as

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

and that of Y is given as

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

6) The **conditional probability density function** of Y given X = x is defined as

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

and that of X given Y = y is defined as

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

7) The **conditional distribution function** of Y given X = x is defined as

$$F(y \mid x) = \int_{-\infty}^{y} f(y \mid x)\,dy, \text{ for all x such that } f(x) > 0;$$

and that of X given Y = y is defined as

$$F(x \mid y) = \int_{-\infty}^{x} f(x \mid y)\,dx, \text{ for all y such that } f(y) > 0.$$

8) Two random variables are said to be **(stochastically) independent** if and only if their joint probability density function is the product of their marginal density functions.

7.8 SOLUTIONS/ANSWERS

E1) As $f(x, y)$ is the joint probability density function,

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

$$\Rightarrow \int_0^4\int_1^5 kxy\, dy\, dx = 1 \;\Rightarrow\; k\int_0^4\left[\int_1^5 xy\, dy\right]dx = 1$$

$$\Rightarrow k\int_0^4\left[x\frac{y^2}{2}\right]_1^5 dx = 1 \Rightarrow\; k\int_0^4 12x\, dx = 1$$

$$\Rightarrow 12k\left[\frac{x^2}{2}\right]_0^4 = 1 \;\Rightarrow\; 96\, k = 1$$

$$\Rightarrow k = \frac{1}{96}$$

E2) i) Marginal density function of Y is given by

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx \;= \int_y^1 2dx$$

[As x is involved in both the given ranges, i.e. $0 < x < 1$ and $0 < y < x$; therefore, here we will combine both these intervals and hence have $0 < y < x < 1$. $\therefore$ x takes the values from y to 1]

$$= [2x]_y^1 = 2 - 2y$$

$= 2 – 2y$

$= 2(1– y), 0 < y < 1$

Marginal density function of X is given by

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

$$= \int_0^x 2dy \qquad [\because\ 0 < \underbrace{y < x}\ < 1]$$

$$= 2[y]_0^x$$

$$= 2x,\ 0 < x < 1.$$

ii) Conditional density function of Y given X($0 < X < 1$) is

$$f(y\,|\,x) = \frac{f(x, y)}{f(x)} = \frac{2}{2x} = \frac{1}{x};\ 0 < y < x$$

Conditional density function of X and given Y($0 < Y < 1$) is

$$f(x\,|\,y) = \frac{f(x, y)}{f(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y},\ y < x < 1$$

iii) $f(x, y) = 2,$

$f(x)f(y) = 2(2x)(1-y)$

As $f(x, y) \neq f(x)f(y),$

$\therefore$ X and Y are not independent.

E3) (i) $P[X<1, Y<3] = \int_{-\infty}^{1}\int_{-\infty}^{3} f(x, y)\,dy\,dx$

$$= \int_{0}^{1}\int_{2}^{3}\frac{1}{8}(6-x-y)\,dy\,dx$$

$$= \int_{0}^{1}\left[\frac{1}{8}\left(6y - xy - \frac{y^2}{2}\right)\right]_2^3 dx$$

$$= \frac{1}{8}\int_{0}^{1}\left[\left\{6(3) - x(3) - \frac{9}{2}\right\} - \{12 - 2x - 2\}\right]dx$$

$$= \frac{1}{8}\int_{0}^{1}\left[\left(18 - 3x - \frac{9}{2}\right) - (10 - 2x)\right]dx$$

$$= \frac{1}{8}\int_{0}^{1}\left(\frac{7}{2} - x\right)dx \;\; = \frac{1}{8}\left[\frac{7}{2}x - \frac{x^2}{2}\right]_0^1 \;\; = \frac{1}{8}\left[\frac{7}{2} - \frac{1}{2}\right] = \frac{3}{8}$$

ii) $P[X<1 \mid Y<3] = \dfrac{P[X<1, Y<3]}{P[Y<3]}$

where $P(Y<3) = \int_{0}^{2}\int_{2}^{3}\frac{1}{8}(6-x-y)\,dy\,dx$

$$= \frac{1}{8}\int_{0}^{2}\left[6y - xy - \frac{y^2}{2}\right]_2^3 dx\,.$$

$$= \frac{1}{8}\int_{0}^{2}\left[\left\{18 - 3x - \frac{9}{2}\right\} - \{12 - 2x - 2\}\right]dx$$

$$= \frac{1}{8}\int_{0}^{2}\left[\left(18 - 3x - \frac{9}{2}\right) - (10 - 2x)\right]dx$$

$$= \frac{1}{8}\int_{0}^{2}\left(\frac{7}{2} - x\right)dx$$

$$= \frac{1}{8}\left[\frac{7}{2}x - \frac{x^2}{2}\right]_0^2$$

$$= \frac{1}{8}\left[7 - \frac{4}{2} - 0\right]$$

$$= \frac{5}{8}$$

$$\therefore P\left[X < 1 \middle| Y < 3\right] = \frac{3/8}{5/8} \quad \left[\begin{array}{l}\because \text{value of numerator is} \\ \text{already calculated in part(i)}\end{array}\right]$$

$$= \frac{3}{5}$$

UNIT 8 MATHEMATICAL EXPECTATION

Structure

8.1 INTRODUCTION

In Units 1 to 4 of this course, you have studied probabilities of different events in various situations. Concept of univariable random variable has been introduced in Unit 5 whereas that of bivariate random variable in Units 6 and 7. Before studying the present unit, we advice you to go through the above units.

You have studied the methods of finding mean, variance and other measures in context of frequency distributions in MST-002 (Descriptive Statistics). Here, in this unit we will discuss mean, variance and other measures in context of probability distributions of random variables. Mean or Average value of a random variable taken over all its possible values is called the expected value or the expectation of the random variable. In the present unit, we discuss the expectations of random variables and their properties.

In Secs. 8.2, 8.3 and 8.4, we deal with expectation and its properties. Addition and multiplication laws of expectation have been discussed in Sec. 8.5.

Objectives

After studying this unit, you would be able to:

- find the expected values of random variables;
- establish the properties of expectation;
- obtain various measures for probability distributions; and
- apply laws of addition and multiplication of expectation at appropriate situations.

8.2 EXPECTATION OF A RANDOM VARIABLE

In Unit 1 of MST-002, you have studied that the mean for a frequency distribution of a variable X is defined as

$$\text{Mean} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}.$$

If the frequency distribution of the variable X is given as

$$\begin{array}{llll} x: & x_1 & x_2 & x_3 \ldots x_n \\ f: & f_1 & f_2 & f_3 \ldots f_n \end{array}$$

The above formula of finding mean may be written as

$$\text{Mean} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_n x_n}{\sum_{i=1}^{n} f_i}$$

$$= \frac{x_1 f_1}{\sum_{i=1}^{n} f_i} + \frac{x_2 f_2}{\sum_{i=1}^{n} f_i} + \ldots + \frac{x_n f_n}{\sum_{i=1}^{n} f_i}$$

$$= x_1 \left(\frac{f_i}{\sum_{i=1}^{n} f_i} \right) + x_2 \left(\frac{f_2}{\sum_{i=1}^{n} f_i} \right) + \ldots + x_n \left(\frac{f_n}{\sum_{i=1}^{n} f_i} \right)$$

Notice that $\frac{f_1}{\sum_{i=1}^{n} f_i}, \frac{f_2}{\sum_{i=1}^{n} f_i}, \ldots, \frac{f_n}{\sum_{i=1}^{n} f_i}$ are, in fact, the relative frequencies or the proportion of individuals corresponding to the values $x_1, x_2, \ldots, x_n$ respectively of variable X and hence can be replaced by probabilities. [See Unit 2 of this course]

Let us now define a similar measure for the probability distribution of a random variable X which assumes the values say $x_1, x_2, \ldots, x_n$ with their associated probabilities $p_1, p_2, \ldots, p_n$. This measure is known as expected value of X and in the similar way is given as

$x_1(p_1) + x_2(p_2) + \ldots + x_n(p_n) = \sum_{i=1}^{n} x_i p_i$ with only difference is that the role of relative frequencies has now been taken over by the probabilities. The expected value of X is written as E(X).

The above aspect can be viewed in the following way also:

Mean of a frequency distribution of X is $\dfrac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$, similarly mean of a

probability distribution of r.v. X is $\dfrac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i}$.

Now, as we know that $\sum_{i=1}^{n} p_i = 1$ for a probability distribution, therefore

the mean of the probability distribution becomes $\sum_{i=1}^{n} x_i p_i$.

$\therefore$ Expected value of a random variable X is $E(X) = \sum_{i=1}^{n} x_i p_i$.

The above formula for finding the expected value of a random variable X is used only if X is a discrete random variable which takes the values $x_1, x_2, ..., x_n$ with probability mass function

$p(x_i) = P[X = x_i]$, $i = 1, 2, ..., n.$

But, if X is a continuous random variable having the probability density function $f(x)$, then in place of summation we will use integration and in this case, the expected value of X is defined as

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

The expectation, as defined above, agrees with the logical/theoretical argument also as is illustrated in the following example.

Suppose, a fair coin is tossed twice, then answer to the question, "How many heads do we expect theoretically/logically in two tosses?" is obviously 1 as the coin is unbiased and hence we will undoubtedly expect one head in two tosses. Expectation actually means "what we get on an average"? Now, let us obtain the expected value of the above question using the formula.

Let X be the number of heads in two tosses of the coin and we are to obtain E(X), i.e. expected number of heads. As X is the number of heads in two tosses of the coin, therefore X can take the values 0, 1, 2 and its probability distribution is given as

X:	0	1	2
$p(x)$:	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

[Refer Unit 5 of MST-003]

$$\therefore\ E(X) = \sum_{i=1}^{3} x_i p_i$$

$$= x_1 p_1 + x_2 p_2 + x_3 p_3$$

$$= (0)\left(\frac{1}{4}\right)+(1)\left(\frac{1}{2}\right)+(2)\left(\frac{1}{4}\right)= 0+\frac{1}{2}+\frac{1}{2}=1$$

So, we get the same answer, i.e. 1 using the formula also.

So, expectation of a random variable is nothing but the average (mean) taken over all the possible values of the random variable or it is the value which we get on an average when a random experiment is performed repeatedly.

Remark 1: Sometimes summations and integrals as considered in the above definitions may not be convergent and hence expectations in such cases do not exist. But we will deal only those summations (series) and integrals which are convergent as the topic regarding checking the convergence of series or integrals is out of the scope of this course. You need not to bother as to whether the series or integral is convergent or not, i.e. as to whether the expectation exists or not as we are dealing with only those expectations which exist.

Example 1: If it rains, a rain coat dealer can earn Rs 500 per day. If it is a dry day, he can lose Rs 100 per day. What is his expectation, if the probability of rain is 0.4?

Solution: Let X be the amount earned on a day by the dealer. Therefore, X can take the values Rs 500, $-$ Rs 100 ($\because$ loss of Rs 100 is equivalent to negative of the earning of Rs100).

$\therefore$ Probability distribution of X is given as

	Rainy Day	Dry day
$X(\text{in Rs.})$:	500	-100
$p(x)$:	0.4	0.6

Hence, the expectation of the amount earned by him is

$$E(X)=\sum_{i=1}^{2} x_i p_i = x_1p_1 + x_2p_2$$

$$= (500)(0.4)+(-100)(0.6) = 200-60 = 140$$

Thus, his expectation is Rs 140, i.e. on an overage he earns Rs 140 per day.

Example 2: A player tosses two unbiased coins. He wins Rs 5 if 2 heads appear, Rs 2 if one head appears and Rs1 if no head appears. Find the expected value of the amount won by him.

Solution: In tossing two unbiased coins, the sample space, is

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

$$\therefore\ P[2\,\text{heads}]=\frac{1}{4},\ \ P(\text{one head})=\frac{2}{4},\ \ P(\text{no head})=\frac{1}{4}.$$

Let X be the amount in rupees won by him

$\therefore$ X can take the values 5, 2 and 1 with

$P[X=5]=P(2\text{heads})=\frac{1}{4},$

$P[X=2]=P[1\text{Head}]=\frac{2}{4},$ and

$P[X=1]=P[\text{no Head}]=\frac{1}{4}.$

$\therefore$ Probability distribution of X is

$X:$	5	2	1
$p(x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Expected value of X is given as

$$E(X)=\sum_{i=1}^{3} x_i p_i = x_1p_1 + x_2p_2 + x_3p_3$$

$$= 5\left(\frac{1}{4}\right)+2\left(\frac{2}{4}\right)+1\left(\frac{1}{4}\right) = \frac{5}{4}+\frac{4}{4}+\frac{1}{4}=\frac{10}{4}=2.5.$$

Thus, the expected value of amount won by him is Rs 2.5.

Example 3: Find the expectation of the number on an unbiased die when thrown.

Solution: Let X be a random variable representing the number on a die when thrown.

$\therefore$ X can take the values 1, 2, 3, 4, 5, 6 with

$P[X=1]=P[X=2]=P[X=3]=P[X=4]=P[X=5]=P[X=6]=\frac{1}{6}.$

Thus, the probability distribution of X is given by

$X:$	1	2	3	4	5	6
$p(x):$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Hence, the expectation of number on the die when thrown is

$$E(X)=\sum_{i=1}^{6} x_i p_i = 1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6} = \frac{21}{6}=\frac{7}{2}$$

Example 4: Two cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the expected value for the number of aces.

Solution: Let A_1, A_2 be the events of getting ace in first and second draws, respectively. Let X be the number of aces drawn. Thus, X can take the values 0, 1, 2 with

$$P[X=0]=P[\text{no ace}]=P\left[\overline{A}_1 \cap \overline{A}_2\right]$$

$$= P\left[\overline{A}_1\right]P\left[\overline{A}_2\right] \quad \left[\begin{array}{l}\because \text{cards are drawn with replacement} \\ \text{and hence the events are independent}\end{array}\right]$$

$$= \frac{48}{52}\times\frac{48}{52} = \frac{12}{13}\times\frac{12}{13} = \frac{144}{169},$$

$$P[X=1] = [\text{one Ace and one other card}]$$

$$= P\left[\left(A_1 \cap \bar{A}_2\right) \cup \left(\bar{A}_1 \cap A_2\right)\right]$$

$$= P\left[A_1 \cap \bar{A}_2\right] + P\left[\bar{A}_1 \cap A_2\right] \quad \begin{bmatrix}\text{By Addition theorem of probability} \\ \text{for mutually exclusive events}\end{bmatrix}$$

$$= P[A_1]P\left[\bar{A}_2\right] + P\left[\bar{A}_1\right]P[A_2] \quad \begin{bmatrix}\text{By multiplication theorem of} \\ \text{probability for independent events}\end{bmatrix}$$

$$= \frac{4}{52}\times\frac{48}{52} + \frac{48}{52}\times\frac{4}{52} = \frac{1}{13}\times\frac{12}{13} + \frac{12}{13}\times\frac{1}{13} = \frac{24}{169}, \text{ and}$$

$$P[X=2] = P[\text{both aces}] = P[A_1 \cap A_2]$$

$$= P[A_1]P[A_2] = \frac{4}{52}\times\frac{4}{52} = \frac{1}{169}.$$

Hence, the probability distribution of random variable X is

$X:$	0	1	2
$p(x):$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

$\therefore$ The expected value of X is given by

$$E(X) = \sum_{i=1}^{3} x_i p_i = 0\times\frac{144}{169} + 1\times\frac{24}{169} + 2\times\frac{1}{169} = \frac{26}{169} = \frac{2}{13}$$

Example 5: For a continuous distribution whose probability density function is given by:

$$f(x) = \frac{3x}{4}(2-x),\ 0 \le x \le 2, \text{ find the expected value of X.}$$

Solution: Expected value of a continuous random variable X is given by

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^2 x\,\frac{3x}{4}(2-x)\,dx = \frac{3}{4}\int_0^2 x^2(2-x)\,dx$$

$$= \frac{3}{4}\int_0^2 \left(2x^2 - x^3\right)dx = \frac{3}{4}\left[2\frac{x^3}{3} - \frac{x^4}{4}\right]_0^2 = \frac{3}{4}\left[2\frac{(2)^3}{3} - \frac{(2)^4}{4} - 0\right]$$

$$= \frac{3}{4}\left[\frac{16}{3} - \frac{16}{4}\right] = \frac{3}{4}\times\frac{16}{12} = 1$$

Now, you can try the following exercises.

E1) You toss a fair coin. If the outcome is head, you win Rs 100; if the outcome is tail, you win nothing. What is the expected amount won by you?

E2) A fair coin is tossed until a tail appears. What is the expectation of number of tosses?

E3) The distribution of a continuous random variable X is defined by

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

Obtain the expected value of X.

Let us now discuss some properties of expectation in the next section.

8.3 PROPERTIES OF EXPECTATION OF ONE-DIMENSIONAL RANDOM VARIABLE

Properties of mathematical expectation of a random variable X are:

1. E(k) = k, where k is a constant
2. E(kX) = kE(X), k being a constant.
3. E(aX + b) = aE(X) + b, where a and b are constants

Proof:

Discrete case:

Let X be a discrete r.v. which takes the values $x_1, x_2, x_3, \ldots$ with respective probabilities $p_1, p_2, p_3, \ldots$

1. $E(k) = \sum_i k\, p_i$ [By definition of the expectation]

$$= k \sum_i p_i$$

$$= k(1) = k \qquad \left[\begin{array}{l}\because \text{sum of probabilities of all the} \\ \text{possible value of r.v. is 1}\end{array}\right]$$

2. $E(kX) = \sum_i (kx_i)\, p_i$ [By def.]

$$= k \sum_i x_i\, p_i$$

$$= k\, E(X)$$

3. $E(a\,X + b) = \sum_i (ax_i + b)\, p_i$ [By def.]

$$= \sum_i (ax_i p_i + bp_i) = \sum_i ax_i p_i + \sum_i bp_i = a \sum_i x_i p_i + b \sum_i p_i$$

$$= aE(X) + b(1) = aE(X) + b$$

Continuous Case:

Let X be continuous random variable having f(x) as its probability density function. Thus,

1. $E(k) = \int_{-\infty}^{\infty} kf(x)dx$ [By def.]

$= k\int_{-\infty}^{\infty} f(x)dx$

$= k(1) = k$ $\left[\because \text{integral of the p.d.f. over the entire range is 1}\right]$

2. $E(kX) = \int_{-\infty}^{\infty} (kx)f(x)dx$ [By def.]

$= k\int_{-\infty}^{\infty} xf(x)dx = kE(X)$

3. $E(aX+b) = \int_{-\infty}^{\infty} (ax+b)f(x)dx = \int_{-\infty}^{\infty} (ax)f(x)dx + \int_{-\infty}^{\infty} bf(x)dx$

$= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx = aE(X)+b(1) = aE(X)+b$

Example 6: Given the following probability distribution:

X	–2	–1	0	1	2
$p(x)$	0.15	0.30	0	0.30	0.25

Find i) E(X)

ii) E(2X + 3)

iii) $E(X^2)$

iv) E(4X – 5)

Solution

i) $E(X) = \sum_{i=1}^{5} x_i p_i = x_1p_1 + x_2p_2 + x_3p_3 + x_4p_4 + x_5p_5$

$= (-2)(0.15) + (-1)(0.30) + (0)(0) + (1)(0.30) + (2)(0.25)$

$= -0.3 - 0.3 + 0 + 0.3 + 0.5 = 0.2$

ii) E(2X + 3) = $2E(X) + 3$ [Using property 3 of this section]

= 2(0.2) + 3 [Using solution (i) of the question]

= 0.4 + 3 = 3.4

iii) $E(X^2) = \sum_{i=1}^{5} x_i^2 p_i$ [By def.]

$$= x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 + x_4^2 p_4 + x_5^2 p_5$$

$$= (-2)^2 (0.15) + (-1)^2 (0.30) + (0)^2 (0) + (1)^2 (0.30) + (2)^2 (0.25)$$

$$= (4)(0.15) + (1)(0.30) + (0) + (1)(0.30) + (4)(0.25)$$

$$= 0.6 + 0.3 + 0 + 0.3 + 1 = 2.2$$

iv) $E(4X - 5) = E[4X + (-5)]$

$$= 4E(X) + (-5) \qquad \text{[Using property 3 of this section]}$$

$$= 4(0.2) - 5$$

$$= 0.8 - 5 = -4.2$$

Here is an exercise for you.

E4) If X is a random variable with mean ‘μ’ and standard deviation ‘σ’, then what is the expectation of $Z = \dfrac{X - \mu}{\sigma}$?

[**Note:** Here Z so defined is called standard random variate.]

Let us now express the moments and other measures for a random variable in terms of expectations in the following section.

8.4 MOMENTS AND OTHER MEASURES IN TERMS OF EXPECTATIONS

Moments

The moments for frequency distribution have already been studied by you in Unit 3 of MST-002. Here, we deal with moments for probability distributions. The r^{th} order moment about any point ‘A’ (say) of variable X already defined in Unit 3 of MST-002 is given by:

$$\mu_r' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^r}{\sum_{i=1}^{n} f_i}$$

So, the r^{th} order moment about any point ‘A’ of a random variable X having probability mass function $P[X = x_i] = p(x_i) = p_i$ is defined as

$$\mu_r' = \frac{\sum_{i=1}^{n} p_i (x_i - A)^r}{\sum_{i=1}^{n} p_i}$$

[Replacing frequencies by probabilities as discussed in Sec. 8.2 of this unit.]

$$= \sum_{i=1}^{n} p_i (x_i - A)^r \qquad \left[\because \sum_{i=1}^{n} p_i = 1\right]$$

The above formula is valid if X is a discrete random variable. But, if X is a continuous random variable having probability density function f(x), then

r^{th} order moment about A is defined as $\mu_r' = \int_{-\infty}^{\infty} (x-A)^r f(x)\,dx.$

So, r^{th} order moment about any point 'A' of a random variable X is defined as

$$\mu_r' = \begin{cases} \sum_i p_i (x_i - A)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x-A)^r f(x)\,dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X - A)^r$$

Similarly, r^{th} order moment about mean (μ) i.e. r^{th} order central moment is defined as

$$\mu_r = \begin{cases} \sum_i p_i (x_i - \mu)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x-\mu)^r f(x)\,dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X-\mu)^r = E[X - E(X)]^r$$

Variance

Variance of a random variable X is second order central moment and is defined as

$$\mu_2 = V(X) = E[X-\mu]^2 = E[X - E(X)]^2$$

Also, we know that

$$V(X) = \mu_2' - (\mu_1')^2$$

where μ_1', μ_2' be the moments about origin.

$\therefore$ We have $V(X) = E(X^2) - [E(X)]^2$

$$\left[\because \mu_1' = E[X-0]^1 = E(X), \text{ and } \mu_2' = E[X-0]^2 = E(X^2)\right]$$

Theorem 8.1: If X is a random variable, then $V(aX+b) = a^2 V(X)$, where a and b are constants.

Proof: $V(aX+b) = E[(aX+b) - E(aX+b)]^2$ [By def. of variance]

$$= E[aX + b - (aE(X) + b)]^2 \quad \text{[Using property 3 of Sec. 8.3]}$$

$$= E[aX + b - aE(X) - b]^2$$

$$= E[a\{X - E(X)\}]^2$$

$$= E\left[a^2 (X - E(X))^2\right]$$

$$= a^2 E\left[X - E(X)\right]^2 \qquad \text{[Using property 2 of section 8.3]}$$

$$= a^2 V(X) \qquad \text{[By definition of Variance]}$$

Cor. (i) $V(aX) = a^2 V(X)$

(ii) V(b) = 0

(iii) V(X + b) = V(X)

Proof: (i) This result is obtained on putting b = 0 in the above theorem.

(ii) This result is obtained on putting a = 0 in the above theorem.

(iii) This result is obtained on putting a = 1 in the above theorem.

Covariance

For a bivariate frequency distribution, you have already studied in Unit 6 of MST-002 that covariance between two variables X and Y is defined as

$$\text{Cov}(X, Y) = \frac{\sum_i f_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i f_i}$$

$\therefore$ For a bivariate probability distribution, Cov (X, Y) is defined as

$$\text{Cov}(X, Y) = \begin{cases} \sum_i p_{ij}(x_i - \bar{x})(y_i - \bar{y}), \text{ if } (X, Y) \text{ is two-dimensional discrete r.v.} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \bar{x})(y - \bar{y}) f(x, y)\,dy\,dx, \text{ if } (X, Y) \text{ is two dimensional continuous r.v.} \end{cases}$$

where $p_{ij} = P\left[X = x_i, Y = y_j\right]$

$$= E(X - \bar{X})(Y - \bar{Y}) \qquad \text{[By definition of expectation]}$$

$$= E\left[(X - E(X))(Y - E(Y))\right] \qquad \begin{bmatrix} \because E(X) = \text{Mean of X i.e. } \bar{X}, \\ E(Y) = \text{Mean of Y i.e. } \bar{Y} \end{bmatrix}$$

On simplifying,

Cov(X, Y) = E(XY) – E(X) E(Y).

Now, if X and Y are independent random variables then, by multiplication theorem,

E(XY) = E(X)E(Y) and hence in this case Cov(X, Y) = 0.

Remark 2:

i) If X and Y are independent random variables, then

V(X + Y) = V(X) +V(Y).

Proof: $V(X+Y) = E\left[(X+Y) - E(X+Y)\right]^2$

$$= E\left[X + Y - E(X) - E(Y)\right]^2$$

$$= E\left[\{X - E(X)\} + \{Y - E(Y)\}\right]^2$$

$$= E\left[\{X - E(X)\}^2 + \{Y - E(Y)\}^2 + 2\{X - E(X)\}\{Y - E(Y)\}\right]$$

$$= E\left[X - E(X)\right]^2 + E\left[Y - E(Y)\right]^2 + 2E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right]$$

$$= V(X) + V(Y) + 2\text{Cov}(X, Y)$$

$$= V(X) + V(Y) + 0 \qquad [\because \text{ X and Y are independent}]$$

$$= V(X) + V(Y)$$

ii) If X and Y are independent random variables, then

$$V(X - Y) = V(X) + V(Y).$$

Proof: This can be proved in the similar manner as done is Remark 2(i) above.

iii) If X and Y are independent random variables, then

$$V(aX + bY) = a^2V(X) + b^2V(Y).$$

Proof: Prove this result yourself proceeding in the similar fashion as in proof of Remark 2(i).

Mean Deviation about Mean

Mean deviation about mean in context of frequency distribution is

$$\frac{\sum_{i=1}^{n} f_i \,|\, x_i - \bar{x} \,|}{\sum_{i=1}^{n} f_i}, \text{ and}$$

therefore, mean deviation about mean in context of probability distribution is

$$\frac{\sum_{i=1}^{n} p_i \,|\, x \ - \text{mean} \,|}{\sum_{i=1}^{n} p_i} = \sum_{i=1}^{n} p_i \,|\, x - \text{mean} \,|$$

$\therefore$ by definition of expectation, we have

M.D. about mean = E|X − Mean|

= E|X – E(X)|

$$= \begin{cases} \sum p_i \,|\, x - \text{Mean} \,| & \text{for discrete r.v} \\ \int_{-\infty}^{\infty} |\, x - \text{Mean} \,|\, f(x)\,dx & \text{for continuous r.v} \end{cases}$$

Note: Other measures as defined for frequency distributions in MST-002 can be defined for probability distributions also and hence can be expressed in terms of the expectations in the manner as the moments; variance and covariance have been defined in this section of the Unit.

Example 7: Considering the probability distribution given in Example 6, obtain

i) V(X)

ii) V(2X + 3).

Solution:

(i) $V(X) = E(X^2) - [E(X)]^2$

$$= 2.2 - (0.2)^2 \qquad \begin{bmatrix} \text{The values have already been obtained} \\ \text{in the solution of Example 6} \end{bmatrix}$$

$$= 2.2 - 0.04 = 2.16$$

(ii) $V(2X + 3) = (2)^2 \, V(X)$ [Using the result of Theorem 8.1]

$$= 4V(X) = 4(2.16) = 8.64$$

Example 8: If X and Y are independent random variables with variances 2 and 3 respectively, find the variance of 3X + 4Y.

Solution: $V(3X + 4Y) = (3)^2\, V(X) + (4)^2\, V(Y)$ [By Remark 3 of Section 8.4]

$$= 9(2) + 16(3) = 18 + 48 = 66$$

Here are two exercises for you:

E5) If X is a random variable with mean μ and standard deviation σ, then find the variance of standard random variable $Z = \dfrac{X - \mu}{\sigma}$.

E6) Suppose that X is a random variable for which E(X) = 10 and V(X) = 25. Find the positive values of a and b such that Y = aX – b has expectation 0 and variance 1.

8.5 ADDITION AND MULTIPLICATION THEOREMS OF EXPECTATION

Now, we are going to deal with the properties of expectation in case of two-dimensional random variable. Two important properties, i.e. addition and multiplication laws of expectation are discussed in the present section.

Addition Theorem of Expectation

Theorem 8.2: If X and Y are random variables, then $E(X+Y)=E(X)+E(Y)$

Proof:

Discrete case:

Let (X, Y) be a discrete two-dimensional random variable which takes up the values (x_i, y_j) with the joint probability mass function
$p_{ij} = P\left[X = x_i \cap Y = y_j\right]$.

Then, the probability distribution of X is given by

$$p_i = p(x_i) = P[X = x_i]$$

$$= P[X = x_i \cap Y = y_1] + P[X = x_i \cap Y = y_2] + \ldots \begin{bmatrix} \because \text{event } X = x_i \text{ can happen with} \\ Y = y_1 \text{ or } Y = y_2 \text{ or } Y = y_3 \text{ or} \ldots \end{bmatrix}$$

$$= p_{i1} + p_{i2} + p_{i3} + \ldots$$

$$= \sum_j p_{ij}$$

Similarly, the probability distribution of Y is given by

$$p_j' = p(y_j) = P\left[Y = y_j\right] = \sum_i p_{ij}$$

$$\therefore\ E(X) = \sum_i x_i p_i,\ E(Y) = \sum_j y_j p_j' \text{ and } E(X+Y) = \sum_i \sum_j (x_i + y_j) p_{ij}$$

Now $E(X+Y) = \sum_i \sum_j (x_i + y_j) p_{ij}$

$$= \sum_i \sum_j x_i p_{ij} + \sum_i \sum_j y_j p_{ij}$$

$$= \sum_i x_i \sum_j p_{ij} + \sum_j y_j \sum_i p_{ij}$$

[$\because$ in the first term of the right hand side, x_i is free from j and hence can be taken outside the summation over j; and in second term of the right hand side, y_j is free from i and hence can be taken outside the summation over i.]

$$\therefore\ E(X+Y) = \sum_i x_i p_i + \sum_j y_j p_j' = E(X) + E(Y)$$

Continuous Case:

Let (X, Y) be a bivariate continuous random variable with probability density function $f(x, y)$. Let $f(x)$ and $f(y)$ be the marginal probability density functions of random variables X and Y respectively.

$\therefore\ E(X)=\int_{-\infty}^{\infty} x\ f(x)dx,\ E(Y)=\int_{-\infty}^{\infty} y\ f(y)dy,$

and $E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,dy\,dx$.

Now, $E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,dy\,dx$

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

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

[$\because$ in the first term of R.H.S., x is free from the integral w.r.t. y and hence can be taken outside this integral. Similarly, in the second term of R.H.S, y is free from the integral w.r.t. x and hence can be taken outside this integral.]

$$=\int_{-\infty}^{\infty}x\,f(x)\,dx+\int_{-\infty}^{\infty}y\,f(y)\,dy\ \left[\begin{array}{l}\text{Refer to the definition of marginal density}\\\text{function given in Unit 7 of this course}\end{array}\right]$$

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

Remark 3: The result can be similarly extended for more than two random variables.

Multiplication Theorem of Expectation

Theorem 8.3: If X and Y are independent random variables, then

E(XY) = E(X) E(Y)

Proof:

Discrete Case:

Let (X, Y) be a two-dimensional discrete random variable which takes up the values (x_i, y_j) with the joint probability mass function $p_{ij}=P\left[X=x_i\cap Y=y_j\right]$. Let p_i and $p_j^{'}$ be the marginal probability mass functions of X and Y respectively.

$\therefore\ E(X)=\sum_i x_i p_i,\ E(Y)=\sum_j y_j p_j^{'},$ and

$$E(XY)=\sum_i\sum_j(x_i y_j)p_{ij}$$

But as X and Y are independent,

$\therefore\ p_{ij}=P\left[X=x_i\cap Y=y_j\right]$

$$= P\left[X = x_i\right]P\left[Y = y_j\right] \quad \left[\begin{array}{l}\because \text{ if events A and B are independent,} \\ \text{then } P(A \cap B) = P(A)P(B)\end{array}\right]$$

$$= p_i p_j^{'}$$

Hence, $E(XY) = \sum_i \sum_j \left(x_i y_j\right) p_i p_j^{'}$

$$= \sum_i \sum_j x_i y_j p_i p_j^{'}$$

$$= \sum_i \sum_j \left(x_i p_i y_j p_j^{'}\right)$$

$$= \sum_i x_i p_i \sum_j y_j p_j^{'} \quad \left[\begin{array}{l}\because \; x_i p_i \text{ is free from j and hence can be} \\ \text{taken outside the summation over j}\end{array}\right]$$

$$= E(X)\, E(Y)$$

<u>Continuous Case:</u>

Let (X, Y) be a bivariate continuous random variable with probability density function f(x, y). Let f(x) and f(y) be the marginal probability density function of random variables X and Y respectively.

$$\therefore \; E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx, \; E(Y) = \int_{-\infty}^{\infty} y\, f(y)\, dy,$$

and $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x, y)\, dy\, dx$.

Now $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x, y)\, dy\, dx$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x)\, f(y)\, dy\, dx \quad \left[\begin{array}{l}\because \text{ X and Y are independent, } f(x,y) = f(x)f(y) \\ \text{(see Unit 7 of this course)}\end{array}\right]$$

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

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

$$= E(X)E(Y)$$

Remark 4: The result can be similarly extended for more than two random variables.

Example 8: Two unbiased dice are thrown. Find the expected value of the sum of number of points on them.

Solution: Let X be the number obtained on the first die and Y be the number obtained on the second die, then

$E(X)=\frac{7}{2}$ and $E(Y)=\frac{7}{2}$ [See Example 3 given in Section 8.2]

$\therefore$ The required expected value $= E(X + Y)$

$$= E(X) + E(Y) \quad \begin{bmatrix}\text{Using addition theorem} \\ \text{of expectation}\end{bmatrix}$$

$$= \frac{7}{2}+\frac{7}{2} = 7$$

Remark 5: This example can also be done considering one random variable only as follows:

Let X be the random variable denoting "the sum of numbers of points on the dice", then the probability distribution in this case is

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

and hence $E(X) = 2\times\frac{1}{36}+3\times\frac{2}{36}+...+12\times\frac{1}{36} = 7$

Example 9: Two cards are drawn one by one with replacement from 8 cards numbered from 1 to 8. Find the expectation of the product of the numbers on the drawn cards.

Solution: Let X be the number on the first card and Y be the number on the second card. Then probability distribution of X is

X	1	2	3	4	5	6	7	8
$p(x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

and the probability distribution of Y is

Y	1	2	3	4	5	6	7	8
$p(y)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$\therefore \quad E(X)=E(Y)=1\times\frac{1}{8}+2\times\frac{1}{8}+...+8\times\frac{1}{8}$$

$$=\frac{1}{8}(1+2+3+4+5+6+7+8)=\frac{1}{8}(36) = \frac{9}{2}$$

Thus, the required expected value is

$E(XY)=E(X)E(Y)$ [Using multiplication theorem of expectation]

$$=\frac{9}{2}\times\frac{9}{2}=\frac{81}{4}.$$

Expectation of Linear Combination of Random Variables

Theorem 8.4: Let $X_1, X_2, ..., X_n$ be any n random variables and if $a_1, a_2, ..., a_n$ are any n constants, then

$$E\left(a_1X_1 + a_2X_2 + ... + a_nX_n\right) = a_1E\left(X_1\right) + a_2E\left(X_2\right) + ... + a_nE\left(X_n\right)$$

[**Note :** Here $a_1X_1 + a_2X_2 + ... + a_nX_n$ is a linear combination of $X_1, X_2, ..., X_n$]

Proof: Using the addition theorem of expectation, we have

$$E\left(a_1X_1 + a_2X_2 + ... + a_nX_n\right) = E\left(a_1X_1\right) + E\left(a_2X_2\right) + ... + E\left(a_nX_n\right)$$

$$= a_1E\left(X_1\right) + a_2E\left(X_2\right) + ... + a_nE\left(X_n\right).$$

[Using second property of Section 8.3 of the unit]

Now, you can try the following exercises.

E7) Two cards are drawn one by one with replacement from ten cards numbered 1 to 10. Find the expectation of the sum of points on two cards.

E8) Find the expectation of the product of number of points on two dice.

Now before ending this unit, let's summarize what we have covered in it.

8.6 SUMMARY

The following main points have been covered in this unit:

1) Expected value of a random variable X is defined as

$$E\left(X\right) = \sum_{i=1}^{n} x_i p_i \text{ , if X is a discrete random variable}$$

$$= \int_{-\infty}^{\infty} x f\left(x\right) dx \text{ , if X is a continuous random variable.}$$

2) Important properties of expectation are:

   i) E(k) = k, where k is a constant.

   ii) E(kX) = kE(X), k being a constant.

   iii) E(aX + b) = aE(X) + b, where a and b are constants

   iv) Addition theorem of Expectation is stated as:

   If X and Y are random variables, then $E\left(X+Y\right) = E\left(X\right) + E\left(Y\right)$.

   v) Multiplication theorem of Expectation is stated as:

   If X and Y are independent random variables, then

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

vi) If $X_1, X_2, ..., X_n$ be any n random variables and if $a_1, a_2, ..., a_n$ are any n constants, then

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n).$$

3) Moments and other measures in terms of expectation are given as:

i) r^{th} order moment about any point is given as

$$\mu_r' = \begin{cases} \sum_i p_i (x_i - A)^r, & \text{if X is a discrete r.v.} \\ \int_{-\infty}^{\infty} (x-A)^r f(x)dx, & \text{if X is a continous r.v} \end{cases}$$

$$= E(X - A)^r$$

ii) Variance of a random variable X is given as

$$V(X) = E[X-\mu]^2 = E[X - E(X)]^2$$

iii) $\text{Cov}(X,Y) = \begin{cases} \sum_i p_i (x_i - \bar{x})(y_i - \bar{y}), & \text{if } (X, Y) \text{ is discrete r.v.} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})(y-\bar{y}) f(x,y)dydx, & \text{if } (X, Y) \text{ is continuous r.v.} \end{cases}$

$$= E[(X - E(X))(Y - E(Y))]$$

$$= E(XY) - E(X)\,E(Y).$$

iv) M.D. about mean $= E|X - E(X)|$

$$= \begin{cases} \sum p_i \,|x - \text{Mean}| & \text{for discrete r.v} \\ \int_{-\infty}^{\infty} |x - \text{Mean}| f(x)dx & \text{for continuous r.v} \end{cases}$$

If you want to see what our solutions to the exercises in the unit are, we have given them in the following section.

8.7 SOLUTIONS/ANSWERS

E1) Let X be the amount (in rupees) won by you.

$\therefore$ X can take the values 100, 0 with P[X = 100] = P[Head] = $\frac{1}{2}$, and

P[X = 0] = P[Tail] = $\frac{1}{2}$.

$\therefore$ probability distribution of X is

X :	100	0
p(x)	$\frac{1}{2}$	$\frac{1}{2}$

and hence the expected amount won by you is

$$E(X) = 100 \times \frac{1}{2} + 0 \times \frac{1}{2} = 50.$$

E2) Let X be the number of tosses till tail turns up.

$\therefore$ X can take values 1, 2, 3, 4… with

P[X = 1] = P[Tail in the first toss] = $\frac{1}{2}$

P[X = 2] = P[Head in the first and tail in the second toss] = $\frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^2$,

P[X = 3] = P[HHT] = $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^3$, and so on.

$\therefore$ Probability distribution of X is

X :	1	2	3	4	5...
$p(x)$	$\frac{1}{2}$	$\left(\frac{1}{2}\right)^2$	$\left(\frac{1}{2}\right)^3$	$\left(\frac{1}{2}\right)^4$	$\left(\frac{1}{2}\right)^5$ ...

and hence

$$E(X) = 1 \times \frac{1}{2} + 2 \times \left(\frac{1}{2}\right)^2 + 3 \times \left(\frac{1}{2}\right)^3 + 4 \times \left(\frac{1}{2}\right)^4 + ... \qquad \text{… (1)}$$

Multiplying both sides by $\frac{1}{2}$, we get

$$\frac{1}{2}E(X) = \left(\frac{1}{2}\right)^2 + 2 \times \left(\frac{1}{2}\right)^3 + 3 \times \left(\frac{1}{2}\right)^4 + 4 \times \left(\frac{1}{2}\right)^5 + ...$$

$$\Rightarrow \frac{1}{2}E(X) = \left(\frac{1}{2}\right)^2 + 2 \times \left(\frac{1}{2}\right)^3 + 3 \times \left(\frac{1}{2}\right)^4 + ... \qquad \text{… (2)}$$

[Shifting the position one step towards right so that we get the terms having same power at the same positions as that in (1)]

Now, subtracting (2) from (1), we have

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

$$\Rightarrow \frac{1}{2}E(X) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4$$

$$\Rightarrow E(X) = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + ...$$

(Which is an infinite G.P. with first term a = 1 and common ratio $r = \frac{1}{2}$)

$$= \frac{1}{1-\frac{1}{2}} \qquad \left[\because S_\infty = \frac{a}{1-r} \text{ (see Unit 3 of course MST - 001)}\right]$$

$$= \frac{1}{\frac{1}{2}} = 2.$$

E3) $E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$

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

$$= \int_{-\infty}^{0} x(0)\, dx + \int_{0}^{1} x\left(x^3\right) dx + \int_{1}^{2} x(2-x)^3\, dx + \int_{2}^{\infty} x(0)\, dx$$

$$= 0 + \int_{0}^{1} x^4 dx + \int_{1}^{2} x\left[8 - x^3 - 6x(2-x)\right] dx + 0$$

$$= \int_{0}^{1} x^4 dx + \int_{1}^{2} \left(8x - x^4 - 12x^2 + 6x^3\right) dx$$

$$= \left[\frac{x^5}{5}\right]_0^1 + \left[8\frac{x^2}{2} - \frac{x^5}{5} - 12\frac{x^3}{3} + 6\frac{x^4}{4}\right]_1^2$$

$$= \frac{1}{5} + \left[\left\{\frac{8(2)^2}{2} - \frac{(2)^5}{5} - \frac{12(2)^3}{3} + \frac{6(2)^4}{4}\right\} - \left\{\frac{8(1)^2}{2} - \frac{(1)^5}{5} - \frac{12(1)^3}{3} + \frac{6(1)^4}{4}\right\}\right]$$

$$= \frac{1}{5} + \left[\left\{16 - \frac{32}{5} - 32 + 24\right\} - \left\{4 - \frac{1}{5} - 4 + \frac{3}{2}\right\}\right]$$

$$= \frac{1}{5} + \left[\frac{8}{5} - \frac{13}{10}\right] = \frac{1}{5} + \frac{3}{10} = \frac{1}{2}.$$

E4) As X is a random variable with mean μ,

$\therefore$ E(X) = μ ... (1)

$$\left[\begin{array}{l}\because \text{expectation is nothing but simply the average taken over all} \\ \text{the possible values of random variable as defined in Sec. 8.2}\end{array}\right]$$

Now, $E(Z) = E\left(\frac{X-\mu}{\sigma}\right)$

$$= E\left[\frac{1}{\sigma}(X-\mu)\right]$$

$$= \frac{1}{\sigma}E[X-\mu] \qquad \text{[Using Property 2 of Sec. 8.3]}$$

$$= \frac{1}{\sigma}\left[E(X)-\mu\right] \qquad \text{[Using Property 3 of Sec. 8.3]}$$

$$= \frac{1}{\sigma}[\mu-\mu] \qquad \text{[Using (1)]}$$

$$= 0$$

Note: Mean of standard random variable is zero.

E5) Variance of standard random variable Z = $\frac{X-\mu}{\sigma}$ is given as

$$V(Z) = V\left(\frac{X-\mu}{\sigma}\right) = V\left(\frac{X}{\sigma}-\frac{\mu}{\sigma}\right)$$

$$= V\left[\frac{1}{\sigma}X+\left(-\frac{\mu}{\sigma}\right)\right]$$

$$= \left(\frac{1}{\sigma}\right)^2 V(X) \qquad \left[\begin{array}{l}\text{Using the result of the Theorem 8.1}\\ \text{of Sec. 8.5 of this unit}\end{array}\right]$$

$$= \frac{1}{\sigma^2}V(X)$$

$$= \frac{1}{\sigma^2}\left(\sigma^2\right) = 1 \qquad \left[\begin{array}{l}\because \text{it is given that the standard deviation}\\ \text{of X is and hence its variance is } \sigma^2\end{array}\right]$$

Note: The mean of standard random variate is '0' [See **(E4)**] and its variance is 1.

E6) Given that $E(Y) = 0 \Rightarrow E(a\,X - b) = 0 \Rightarrow a\,E(X) - b = 0$

$\Rightarrow a(10) - b = 0$

$\Rightarrow 10\,a - b = 0$ ... (1)

Also as V(Y) = 1,

hence V(aX – b) = 1

$\Rightarrow a^2V(X) = 1 \Rightarrow a^2(25) = 1 \Rightarrow a^2 = \frac{1}{25}$

$\Rightarrow a = \frac{1}{5}$ [$\because$ a is positive]

$\therefore$ From (1), we have

$$10\left(\frac{1}{5}\right) - b = 0 \Rightarrow 2 - b = 0 \Rightarrow b = 2$$

Hence, $a = \frac{1}{5}$, b = 2.

E7) Let X be the number on the first card and Y be the number on the second card. Then probability distribution of X is:

X	1	2	3	4	5	6	7	8	9	10
$p(x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

and the probability distribution of Y is

X	1	2	3	4	5	6	7	8	9	10
$p(x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

$$\therefore\ E(X) = E(Y) = 1\times\frac{1}{10}+2\times\frac{1}{10}+...+10\times\frac{1}{10}$$

$$=\frac{1}{10}\left[1+2+3+4+5+6+7+8+9+10\right] = \frac{1}{10}(55)=5.5$$

and hence the required expected value is

$$E(X+Y)=E(X)+E(Y)= 5.5 + 5.5 = 11$$

E8) Let X be the number obtained on the first die and Y be the number obtained on the second die.

Then $E(X) = E(Y) = \frac{7}{2}$. [See Example 3 given in Section 8.2]

Hence, the required expected value is

$$E(XY)=E(X)E(Y) \quad \text{[Using multiplication theorem of expectation]}$$

$$=\frac{7}{2}\times\frac{7}{2} = \frac{49}{4}.$$

UNIT 9 BINOMIAL DISTRIBUTION

Structure

9.1 INTRODUCTION

In Unit 5 of the Course, you have studied random variables, their probability functions and distribution functions. In Unit 8 of the Course, you have come to know as to how the expectations and moments of random variables are obtained. In those units, the definitions and properties of general discrete and probability distributions have been discussed.

The present block is devoted to the study of some special discrete distributions and in this list, Bernoulli and Binomial distributions are also included which are being discussed in the present unit of the course.

Sec. 9.2 of this unit defines Bernoulli distribution and its properties. Binomial distribution and its applications are covered in Secs. 9.3 and 9.4 of the unit.

Objectives

Study of the present unit will enable you to:

- define the Bernoulli distribution and to establish its properties;
- define the binomial distribution and establish its properties;
- identify the situations where these distributions are applied;
- know as to how binomial distribution is fitted to the given data; and
- solve various practical problems related to these distributions.

9.2 BERNOULLI DISTRIBUTION AND ITS PROPERTIES

There are experiments where the outcomes can be divided into two categories with reference to presence or absence of a particular attribute or characteristic. A convenient method of representing the two is to designate either of them as success and the other as failure. For example, head coming up in the toss of a fair coin may be treated as a success and tail as failure, or vice-versa. Accordingly, probabilities can be assigned to the success and failure.

Suppose a piece of a product is tested which may be defective (failure) or non-defective (a success). Let p the probability that it found non-defective and q = 1 – p be the probability that it is defective. Let X be a random variable such that it takes value 1 when success occurs and 0 if failure occurs. Therefore,

$P[X=1]=p,$ and

$P[X=0]=q=1-p.$

The above experiment is a Bernoulli trial, the r.v. X defined in the above experiment is a Bernoulli variate and the probability distribution of X as specified above is called the Bernoulli distribution in honour of J. Bernoulli (1654-1705).

Definition

A discrete random variable X is said to follow Bernoulli distribution with parameter p if its probability mass function is given by

$$P[X=x]=\begin{cases} p^x(1-p)^{1-x} & ; x=0,1 \\ 0 & ; \text{elsewhere} \end{cases}$$

i.e. $P[X=1]=p^1(1-p)^{1-1}=p$ [putting x = 1]

and $P[X=0]=p^0(1-p)^{1-0}=1-p$ [putting x = 0]

The Bernoulli probability distribution, in tabular form, is given as

X	0	1
$p(x)$	$1-p$	p

Remark 1: The Bernoulli distribution is useful whenever a random experiment has only two possible outcomes, which may be labelled as success and failure.

Moments of Bernoulli Distribution

The r^{th} moment about origin of a Bernoulli variate X is given as

$$\mu_r' = E(X^r)$$

$$= \sum_{x=0}^{1} x^r p(x) \qquad \text{[See Unit 8 of this course]}$$

$$= (0)^r p(0) + (1)^r p(1)$$

$$= (0)(1-p) + (1)p$$

$$= p$$

$\Rightarrow \mu_1' = p,\ \mu_2' = p,\ \mu_3' = p,\ \mu_4' = p.$

Hence,

Mean = $\mu_1' = p,$

Variance $(\mu_2) = \mu_2' - (\mu_1')^2 = p - p^2 = p(1-p),$

Third order central moment $(\mu_3) = \mu_3' - 3\mu_2'(\mu_1') + 2(\mu_1')^3$

$$= p - 3pp + 2(p)^3$$

$$= p - 3p^2 + 2p^3$$
$$= p\left(2p^2 - 3p + 1\right) = p\left(2p - 1\right)\left(p - 1\right)$$
$$= p\left(1 - p\right)\left(1 - 2p\right)$$

Fourth order central moment $(\mu_4) = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4$

$$= p - 4p.p + 6p\left(p\right)^2 - 3\left(p\right)^4$$
$$= p - 4p^2 + 6p^3 - 3p^4$$
$$= p\left[1 - 4p + 6p^2 - 3p^3\right]$$
$$= p\left(1 - p\right)\left(1 - 3p + 3p^2\right)$$

[**Note:** For relations of central moments in terms of moments about origin, see Unit 3 of MST-002.]

Example 1: Let X be a random variable having Bernoulli distribution with paramete p = 0.4. Find its mean and variance.

Solution:

Mean = p = 0.4,
Variance = p(1 − p) = (0.4) (1 −0.4) = (0.4) (0.6) = 0.24

Single trial is taken into consideration in Bernoulli distribution. But, if trials are performed repeatedly a finite number of times and we are interested in the distribution of the sum of independent Bernoulli trials with the same probability of success in each trial, then we need to study binomial distribution which has been discussed in the next section.

9.3 BINOMIAL PROBABILITY FUNCTION

Here, is this section, we will discuss binomial distribution which was discovered by J. Bernoulli (1654-1705) and was first published eight years after his death i.e. in 1713 and is also known as “Bernoulli distribution for n trials”. Binomial distribution is applicable for a random experiment comprising a finite number (n) of independent Bernoulli trials having the constant probability of success for each trial.

Before defining binomial distribution, let us consider the following example: Suppose a man fires 3 times independently to hit a target. Let p be the probability of hitting the target (success) for each trial and $q\left(=1-p\right)$ be the probability of his failure.

Let S denote the success and F the failure. Let X be the number of successes in 3 trials,

P[X = 0] = Probability that target is not hit at all in any trial

= P [Failure in each of the three trials]

$= P\left(F \cap F \cap F\right)$

$= P\left(F\right).P\left(F\right).P\left(F\right)$ [∵ trials are independent]

$= q.q.q$

$= q^3$

This can be written as

$$\boxed{P[X=0] = {}^{3}C_{0}\, p^{0} q^{3-0}}$$

$\left[\because\ {}^{3}C_{0} = 1,\ p^{0} = 1,\ q^{3-0} = q^{3}\right.$. Recall ${}^{n}C_{x} = \dfrac{\lfloor n}{\lfloor x \lfloor n-x}$ (see Unit 4 of MST-001)]

P[X = 1] = Probability of hitting the target once

= [(Success in the first trial and failure in the second and third trial) or (success in the second trial and failure in the first and third trials) or (success in the third trial and failure in the first two trials)]

$$= P\left[(S \cap F \cap F) \text{ or } (F \cap S \cap F) \text{ or } (F \cap F \cap S)\right]$$

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

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

[$\because$ trials are independent]

$$= p.q.q + q.p.q + q.q.p$$

$$= pq^{2} + pq^{2} + pq^{2}$$

$$= 3pq^{2}$$

This can also be written as

$$\boxed{P[X=1] = {}^{3}C_{1} p^{1} q^{3-1}} \qquad \left[\because\ {}^{3}C_{1} = 3,\ p^{1} = p,\ q^{3-1} = q^{2}\right]$$

P[X = 2] = Probability of hitting the target twice

= P[(Success in each of the first two trials and failure in the third trial) or (Success in first and third trial and failure in the second trial) or (Success in the last two trials and failure in the first trial)]

$$= P\left[(S \cap S \cap F) \cup (S \cap F \cap S) \cup (F \cap S \cap S)\right]$$

$$= P[S \cap S \cap F] + P[S \cap F \cap S] + P[F \cap S \cap S]$$

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

$$= p.p.q + p.q.p + q.p.p$$

$$= 3\,p^{2}q$$

This can also be written as

$$\boxed{P[X=2] = {}^{3}C_{2} p^{2} q^{3-2}} \qquad \left[\because\ {}^{3}C_{2} = 3, q^{3-2} = q\right]$$

P[X = 3] = Probability of hitting the target thrice

= [Success in each of the three trials]

$$= P[S \cap S \cap S]$$

$$= P(S).P(S).P(S)$$

$$= p.p.p$$

$$= p^{3}$$

This can also be written as

$\boxed{P[X=3] = {}^3C_3\, p^3 q^{3-3}}$ $\quad [\because {}^3C_3 = 1, q^{3-3} = 1]$

From the above four enrectangled results, we can write

$P[X=r] = {}^3C_r p^r q^{3-r}$; $r = 0, 1, 2, 3$.

which is the probability of r successes in 3 trials. ${}^3C_r$, here, is the number of ways in which r successes can happen in 3 trials.

The result can be generalized for n trials in the similar fashion and is given as
$P[X=r] = {}^nC_r p^r q^{n-r}$; $r = 0, 1, 2, ..., n$.

This distribution is called the binomial probability distribution. The reason behind giving the name binomial probability distribution for this probability distribution is that the probabilities for x = 0, 1, 2, …, n are the respective probabilities ${}^nC_0 p^0 q^{n-0}$, ${}^nC_1 p^1 q^{n-1}$, ..., ${}^nC_n p^n q^{n-n}$ which are the successive terms of the binomial expansion $(q + p)^n$.

$[\because (q+p)^n = {}^nC_0 q^n p^0 + {}^nC_1 q^{n-1} p^1 + ... + {}^nC_n q^0 p^n]$

Binomial Expansion:

'Bi' means 'Two'. 'Binomial expansion' means 'Expansion of expression having two terms, e.g.

$$(X+Y)^2 = X^2 + 2XY + Y^2 = {}^2C_0 X^2 Y^0 + {}^2C_1 X^{2-1} Y^1 + {}^2C_2 X^{2-2} Y^2,$$

$$(X+Y)^3 = X^3 + 3X^2Y + 3XY^2 + Y^3$$

$$= {}^3C_0 X^3 Y^0 + {}^3C_1 X^{3-1} Y^1 + {}^3C_2 X^{3-2} Y^2 + {}^3C_3 X^{3-3} Y^3$$

So, in general,

$$(X+Y)^n = {}^nC_0 X^n Y^0 + {}^nC_1 X^{n-1} Y^1 + {}^nC_2 X^{n-2} Y^2 + ... + {}^nC_n X^{n-n} Y^n$$

The above discussion leads to the following definition.

Definition:

A discrete random variable X is said to follow binomial distribution with parameters n and p if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X=x] = \begin{cases} {}^nC_x p^x q^{n-x}; & x = 0, 1, 2, ..., n \\ 0; & \text{elsewhere} \end{cases}$$

where, n is the number of independent trials,

x is the number of successes in n trials,

p is the probability of success in each trial, and

q = 1 – p is the probability of failure in each trial.

Remark 2:

i) The binomial distribution is the probability distribution of sum of n independent Bernoulli variates.

ii) If X is binomially distributed r.v. with parameters n and p, then we may write it as X ~ B(n, p).

iii) If X and Y are two binomially distributed independent random variables with parameters (n_1, p) and (n_2, p) respectively then their sum also follows a binomial distribution with parameters $n_1 + n_2$ and p. But, if the probability of success is not same for the two random variables then this property does not hold.

Example 2: An unbiased coin is tossed six times. Find the probability of obtaining

(i) exactly 3 heads

(ii) less than 3 heads

(iii) more than 3 heads

(iv) at most 3 heads

(v) at least 3 heads

(vi) more than 6 heads

Solution: Let p be the probability of getting head (success) in a toss of the coin and n be the number of trials.

$\therefore$ n = 6, p = $\frac{1}{2}$ and hence q = 1 – p = $1-\frac{1}{2}=\frac{1}{2}$.

Let X be the number of successes in n trials,

$\therefore$ by binomial distribution, we have

$$P\left[X=x\right]={}^{n}C_{x}p^{x}q^{n-x}; x=0, 1, 2, ..., n$$

$$={}^{6}C_{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{6-x}; \; x=0, 1, 2, ..., 6$$

$$={}^{6}C_{x}\left(\frac{1}{2}\right)^{6}; x=0, 1, 2, ..., 6.$$

$$=\frac{1}{64}.{}^{6}C_{x}; \; x=0, 1, 2, ..., 6.$$

Therefore,

(i) P[exactly 3 heads] = P [X = 3]

$$=\frac{1}{64}\left({}^{6}C_{3}\right)=\frac{1}{64}\left[\frac{6\times5\times4}{3\times2}\right]=\frac{5}{16}$$

[$\because$ Recall ${}^{n}C_{x}=\frac{\lfloor n}{\lfloor x \lfloor n-x}$ (see Unit 4 of MST- 001)]

(ii) P[less than 3 heads] = P[X < 3]

$$=P\left[X=2 \text{ or } X=1 \text{ or } X=0\right]$$

$$=P\left[X=2\right]+P\left[X=1\right]+P\left[X=0\right]$$

$$=\frac{1}{64}.{}^{6}C_{2}+\frac{1}{64}.{}^{6}C_{1}+\frac{1}{64}.{}^{6}C_{0}$$

$$= \frac{1}{64}\left[{}^6C_2 + {}^6C_1 + {}^6C_0\right] = \frac{1}{64}\left[\frac{6\times5}{2}+6+1\right]$$

$$= \frac{22}{64} = \frac{11}{32}.$$

(iii) P[more than 3 heads] = P[X > 3]

$$= P[X = 4 \text{ or } X = 5 \text{ or } X = 6] \quad \left[\begin{array}{l}\because \text{ in 6 trials one can} \\ \text{have at most 6 heads}\end{array}\right]$$

$$= P[X=4] + P[X=5] + P[X=6]$$

$$= \frac{1}{64}.{}^6C_4 + \frac{1}{64}.{}^6C_5 + \frac{1}{64}.{}^6C_6$$

$$= \frac{1}{64}\left[{}^6C_4 + {}^6C_5 + {}^6C_6\right]$$

$$= \frac{1}{64}\left[\frac{6\times5}{2}+6+1\right] = \frac{22}{64} = \frac{11}{32}.$$

(iv) P[at most 3 heads] = P [3 or less than 3 heads]

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

$$= \frac{1}{64}.{}^6C_3 + \frac{1}{64}.{}^6C_2 + \frac{1}{64}.{}^6C_1 + \frac{1}{64}.{}^6C_0$$

$$= \frac{1}{64}\left[{}^6C_3 + {}^6C_2 + {}^6C_1 + {}^6C_0\right]$$

$$= \frac{1}{64}\left[20+15+6+1\right] = \frac{42}{64} = \frac{21}{32}.$$

(v) P[at least 3 heads] = P[3 or more heads]

$$= P[X=3] + P[X=4] + P[X=5] + P[X=6]$$

or

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

$$\left[\begin{array}{l}\because \text{ sum of probabilities of all possible} \\ \text{ values of a random variable is 1}\end{array}\right]$$

$$= 1 - \left(\frac{11}{32}\right) \quad \left[\begin{array}{l}\text{Already obtained in} \\ \text{part (ii) of this example}\end{array}\right]$$

$$= \frac{21}{32}.$$

(vi) P [more than 6 heads] = P [7 or more heads]

$$= P\,[\text{an impossible event}] \quad \left[\begin{array}{l}\because \text{ in six tosses, it} \\ \text{is impossible to get} \\ \text{more than six heads}\end{array}\right]$$

$$= 0$$

Example 3: The chances of catching cold by workers working in an ice factory during winter are 25%. What is the probability that out of 5 workers 4 or more will catch cold?

Solution: Let catching cold be the success and p be the probability of success for each worker.

$\therefore$ Here, n = 5, p = 0.25, q = 0.75 and by binomial distribution

$$P[X = x] = {}^{n}C_{x}\, p^{x} q^{n-x} \;;\; x = 0, 1, 2, ..., n$$

$$= {}^{5}C_{x}\,(0.25)^{x}(0.75)^{5-x} \;;\; 0, 1, 2, ..., 5$$

Therefore, the required probability = P[X ≥ 4]

$$= p[X = 4 \text{ or } X = 5]$$

$$= P[X = 4] + P[X = 5]$$

$$= {}^{5}C_{4}(0.25)^{4}(0.75)^{1} + {}^{5}C_{5}(0.25)^{5}(0.75)^{0}$$

$$= (5)(0.002930) + (1)(0.000977)$$

$$= 0.014650 + 0.000977$$

$$= 0.015627$$

Example 4: Let X and Y be two independent random variables such that X ~ B(4, 0.7) and Y ~ B(3, 0.7). Find $P[X + Y \le 1]$.

Solution: We know that if X and Y are independent random variables each following binomial distribution with parameters (n_1, p) and (n_2, p), then X + Y ~ B($n_1 + n_2$, p).

Therefore, here X + Y follows binomial distribution with parameters 4 + 3 and 0.7, i.e. 7 and 0.7. So, here, n = 7 and p = 0.7.

Thus, the required probability = $[X + Y \le 1]$

$$= P[X + Y=1] + P[X + Y= 0]$$

$$= {}^{7}C_{1}(0.7)^{1}(0.3)^{6} + {}^{7}C_{0}(0.7)^{0}(0.3)^{7}$$

$$= 7(0.7)(0.000729) + 1(1)(0.0002187)$$

$$= 0.0035721 + 0.0002187$$

$$= 0.0037908$$

Now, we are sure that you can try the following exercises:

E1) The probability of a man hitting a target is $\frac{1}{4}$. He fires 5 times. What is the probability of his hitting the target at least twice?

E2) A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that the dacoit is still alive?

9.4 MOMENTS OF BINOMIAL DISTRIBUTION

The r^{th} order moment about origin of a binomial variate X is given as

$$\mu_r' = E\left(X^r\right) = \sum_{x=0}^{n} x^r.P\left[X = x\right]$$

$$\therefore \quad \mu_1' = E\left(X\right) = \sum_{x=0}^{n} x\ .P\left[X = x\right]$$

$$= \sum_{x=0}^{n} x.\,{}^nC_x\, p^x q^{n-x} \quad \left[\because P\left[X = x\right] = {}^nC_x p^x q^{n-x}\,;\, x = 0, 1, 2, ..., n\right]$$

$$= \sum_{x=1}^{n} x.\,{}^nC_x\, p^x q^{n-x} \quad \begin{bmatrix} \because \text{first term with } x = 0 \text{ will be zero} \\ \text{and hence we may start from } x = 1 \end{bmatrix}$$

$$= \sum_{x=1}^{n} x.\frac{n}{x}.\ {}^{n-1}C_{x-1}\, p^x q^{n-x}$$

$$\begin{bmatrix} \because\ {}^nC_x = \dfrac{\underline{|n}}{\underline{|x}\,\underline{|n-x}} = \dfrac{n\underline{|n-1}}{x\underline{|x-1}\,\underline{|(n-1)-(x-1)}} = \dfrac{n}{x}\,{}^{n-1}C_{x-1}, \\ \left(\text{see Unit 4 of MST} - 001\right) \end{bmatrix}$$

$$= \sum_{x=1}^{n} n.\,{}^{n-1}C_{x-1}\, p^{x-1}.p^1.q^{(n-1)-(x-1)} \quad [n-x = (n-1) - (x-1)]$$

$$= np\sum_{x=1}^{n}\ {}^{n-1}C_{x-1}\, p^{x-1}.q^{(n-1)-(x-1)}$$

$$= np\begin{bmatrix} {}^{n-1}C_0\, p^0\, q^{(n-1)-0} + {}^{n-1}C_1\, p^1\, q^{(n-1)-1} + {}^{n-1}C_2\, p^2 q^{(n-1)-2} + ... \\ + {}^{n-1}C_{n-1}.\,p^{n-1}\, q^{(n-1)-(n-1)} \end{bmatrix}$$

$$= np \times \begin{bmatrix} \text{Sum of probabilities of all possible values of a} \\ \text{binomial variate with parameters } n-1 \text{ and } p \end{bmatrix}$$

$$= np \times 1 \qquad \begin{bmatrix} \because \text{sum of probabilities of all possible} \\ \text{values of a random variable is 1} \end{bmatrix}$$

$$= np.$$

$\therefore$ Mean = First order moment about origin

$= \mu_1'$

$= np.$

$$\boxed{\text{Mean} = np}$$

$$\mu_2' = E\left(X^2\right) = \sum_{x=0}^{n} x^2.P\left[X = x\right] = \sum_{x=0}^{n} x^2.\,{}^nC_x p^x q^{n-x}$$

Here, we will write x^2 as $x\left(x-1\right)+x$ $\quad [\because x\left(x-1\right)+x = x^2 - x + x = x^2]$

This is done because in the following expression, we get $x\left(x-1\right)$ in the denominator:

$$\left[\because\ {}^nC_x = \frac{\underline{|n}}{\underline{|x}\times\underline{|n-x}} = \frac{n(n-1)\underline{|n-2}}{x(x-1)\underline{|x-2}\times\underline{|(n-2)-(x-2)}} = \frac{n(n-1)}{x(x-1)}.\,{}^{n-2}C_{x-2}\right]$$

$$\therefore\ \mu_2' = \sum_{x=0}^{n}\left[x(x-1)+x\right]{}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\,{}^nC_x p^x\, q^{n-x} + \sum_{x=0}^{n} x.\,{}^nC_x p^x q^{n-x}$$

$$= \left[\sum_{x=2}^{n} x(x-1)\,{}^nC_x p^x q^{(n-2)-(x-2)}\right] + \left(\mu_1'\right)$$

$$= \left[\sum_{x=2}^{n} x(x-1).\frac{n(n-1)}{x(x-1)}\ {}^{n-2}C_{x-2} p^x q^{n-x}\right] + \mu_1'$$

$$= \left[\sum_{x=2}^{n} n(n-1)\,{}^{n-2}C_{x-2} p^{x-2}.p^2 q^{(n-2)-(x-2)}\right] + \mu_1'$$

$$= \left[n(n-1)p^2 \sum_{x=2}^{n} {}^{n-2}C_{x-2} p^2 q^{(n-2)-(x-2)}\right] + \mu_1'$$

$$= n(n-1)p^2 \times \left[\begin{array}{c}\text{Sum of probabilities of all possible values of a}\\ \text{binomial variate with parameters } n-2 \text{ and } p\end{array}\right] + \mu_1'$$

$= n(n-1)\,p^2\,(1) + np \qquad \left[\because \mu_1' = np\right]$

$= n^2\,p^2 - np^2 + np$

$\therefore$ Variance $(\mu_2) = \mu_2' - \left(\mu_1'\right)^2$ [See Unit 3 of MST-002]

$= n^2p^2 - n\,p^2 + np - (np)^2$

$= n^2p^2 - np^2 + np - n^2p^2$

$= np - np^2$

$= np(1-p)$

$= npq$

$$\therefore \quad \boxed{\text{Variance} = npq}$$

$$\mu_3' = \sum_{x=0}^{n} x^3.P\left[X = x\right]$$

Here, we will write x^3 as $x(x-1)(x-2) + 3x(x-1) + x$

Let $x^3 = x(x-1)(x-2) + Bx(x-1) + Cx$ Comparing coefficients of x^2, we have $0 = -3 + B \Rightarrow B = 3$ Comparing coeffs of x, we have $0 = 2 - B + C \Rightarrow C = B - 2 = 3 - 2 \Rightarrow C = 1$

$$\therefore \quad \mu_3' = \sum_{x=0}^{n}\left[x(x-1)(x-2)+3x(x-1)+x\right].{}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\,{}^nC_x p^x q^{n-x} + 3\sum_{x=0}^{n} x(x-1)\,{}^nC_x p^x q^{n-x} + \sum_{x=0}^{n} x.\,{}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\frac{n}{x}.\frac{n-1}{x-1}.\frac{n-2}{x-2}\,{}^{n-3}C_{x-3}p^x q^{n-x} + 3\left[n(n-1)p^2\right]+\left[np\right]$$

[The expression within brackets in the second term is the first term of R.H.S. in the derivation of μ_2' and the expression in the third term is μ_1' as already obtained.]

$$\left[\begin{aligned} \because\ {}^nC_x &= \frac{\lfloor n}{\lfloor x \lfloor n-x} = \frac{n(n-1)(n-2)\lfloor n-3}{x(x-1)(x-2)\lfloor x-3 \lfloor (n-3)-(x-3)} \\ &= \frac{n(n-1)(n-2)}{x(x-1)(x-2)}.\,{}^{n-3}C_{x-3} \end{aligned}\right]$$

$$= \sum_{x=3}^{n} n(n-1)(n-2).{}^{n-3}C_{x-3}p^3 p^{x-3} q^{(n-3)-(x-3)} + 3n(n-1)p^2 + np$$

$$= n(n-1)(n-2)p^3\sum_{x=3}^{n} {}^{n-3}C_{x-3}p^{x-3} q^{(n-3)-(x-3)} + 3n(n-1)p^2 + np$$

$$= n(n-1)(n-2)p^3(1) + 3n(n-1)p^2 + np$$

$\therefore$ Third order central moment is given by

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\left(\mu_1'\right)^3 \qquad \text{[See Unit 4 of MST-002]}$$

$$= npq(q-p) \qquad \text{[On simplification]}$$

$$\boxed{\mu_3 = npq(q-p)}$$

$$\mu_4' = \sum_{x=0}^{n} x^4 P(X=x)$$

Writing

$$x^4 = x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x$$

and proceeding in the similar fashion as for $\mu_1', \mu_2', \mu_3',$ we have

$$\mu_4' = n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np$$

and hence

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - \left(\mu_1'\right)^4$$

$$\mu_4 = npq\left[1+3(n-2)pq\right] \qquad \text{[On simplification]}$$

Now, recall the measures of skewness and kurtosis which you have studied in Unit 4 of MST-002

These measures are given as follows:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left[npq(q-p)\right]^2}{\left[npq\right]^3} = \frac{(q-p)^2}{npq},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\left[1+3(n-2)pq\right]}{\left[npq\right]^2} = 3 + \frac{1-6pq}{npq},$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}}, \text{ and}$$

$$\gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}$$

Remark 3:

(i) As $0 < q < 1$

$\Rightarrow q < 1$

$\Rightarrow npq < np$ [Multiplying both sides by $np > 0$]

$\Rightarrow$ Variance < Mean

Hence, for binomial distribution

Mean > Variance

(ii) As variance of $X \sim B(n, p)$ is npq,

$\therefore$ its standard deviation is $\sqrt{npq}$.

Example 4: For a binomial distribution with $p = \frac{1}{4}$ and $n = 10$, find mean and variance.

Solution: As $p = \frac{1}{4}$, $\therefore\ q = 1 - \frac{1}{4} = \frac{3}{4}$.

Mean = $np = 10 \times \frac{1}{4} = \frac{5}{2}$,

Variance = $npq = 10 \times \frac{1}{4} \times \frac{3}{4} = \frac{15}{8}$.

Example 5: The mean and standard deviation of binomial distribution are 4 and $\frac{2}{\sqrt{3}}$ respectively. Find $P[X \geq 1]$.

Solution: Let $X \sim B(n, p)$, then

Mean = $np = 4$

and variance = $npq = \left(\frac{2}{\sqrt{3}}\right)^2$ [$\because$ S.D. = $\frac{2}{\sqrt{3}}$ and variance is square of S.D.]

Dividing second equation by the first equation, we have

$$\frac{npq}{np} = \frac{\frac{4}{3}}{4}$$

$$\Rightarrow\ q = \frac{1}{3}$$

$$\therefore\ p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$$

Putting p = $\frac{2}{3}$ in the equation of mean, we have

$$n\left(\frac{2}{3}\right) = 4 \Rightarrow n = 6$$

$\therefore$ by binomial distribution,

$$P[X = x] = {}^{n}C_{x}p^{x}q^{n-x}$$

$$= {}^{6}C_{x}\left(\frac{2}{3}\right)^{x}\left(\frac{1}{3}\right)^{6-x} \; ; \; x = 0, 1, 2, \ldots, 6.$$

Thus, the required probability

$$P[X \geq 1] = P[X = 1] + P[X = 2] + P[X = 3] + \ldots + P[X = 6]$$

$$= 1 - P[X = 0]$$

$$= 1 - {}^{6}C_{0}\left(\frac{2}{3}\right)^{0}\left(\frac{1}{3}\right)^{6-0} = 1 - (1)(1)\frac{1}{729} = \frac{728}{729}.$$

Example 6: If X ~ B (n, p). Find p if n = 6 and 9P[X = 4] = P[X = 2].

Solution: As X ~ B(n, p) and n = 6,

$$\therefore \; P[X = x] = {}^{6}C_{x}p^{x}(1-p)^{6-x} \; ; x = 0, 1, 2, \ldots, 6.$$

Now, $9P[X = 4] = P[X = 2]$

$$\Rightarrow 9 \times {}^{6}C_{4} \times p^{4}(1-p)^{6-4} = {}^{6}C_{2} \times p^{2}(1-p)^{4}$$

$$\Rightarrow 9 \times \frac{6\times5}{2} \times p^{4} \times (1-p)^{2} = \frac{6\times5}{2}p^{2}(1-p)^{4}$$

$$\Rightarrow 9p^{2} = (1-p)^{2}$$

$$\Rightarrow 9p^{2} = 1 + p^{2} - 2p$$

$$\Rightarrow 8p^{2} + 2p - 1 = 0$$

$$\Rightarrow 8p^{2} + 4p - 2p - 1 = 0$$

$$\Rightarrow 4p(2p+1) - 1(2p+1) = 0$$

$$\Rightarrow (2p+1)(4p-1) = 0$$

$$\Rightarrow (2p+1) = 0 \text{ or } (4p-1) = 0$$

$$\Rightarrow p = -\frac{1}{2} \text{ or } \frac{1}{4}$$

But $p = -\frac{1}{2}$ rejected [$\because$ probability can never be negative]

Hence, $p = \frac{1}{4}$

Now, you can try the following exercises:

E3) Comment on the following:

The mean of a binomial distribution is 3 and variance is 4.

E4) Find the binomial distribution when sum of mean and variance of 5 trails is 4.8.

E5) The mean of a binomial distribution is 30 and standard deviation is 5. Find the values of

i) n, p and q,

ii) Moment coefficient of skewness, and

iii) Kurtosis.

9.5 FITTING OF BINOMIAL DISTRIBUTION

To fit a binomial distribution, we need the observed data which is obtained from repeated trials of a given experiment. On the basis of the observed data, we find the theoretical (or expected) frequencies corresponding to each value of the binomial variable. Process of finding the probabilities corresponding to each value of the binomial variable becomes easy if we use the recurrence relation for the probabilities of Binomial distribution. So, in this section, we will first establish the recurrence relation for probabilities and then define the binomial frequency distribution followed by process of fitting a binomial distribution.

Recurrence Relation for the Probabilities of Binomial Distribution

You have studied that binomial probability function is

$$p(x) = P[X = x] = {}^nC_x p^x q^{n-x} \qquad \ldots (1)$$

If we replace x by x + 1, we have

$$p(x+1) = {}^nC_{x+1}\, p^{x+1} q^{n-(x+1)} \qquad \ldots (2)$$

Dividing (2) by (1), we have

$$\frac{p(x+1)}{p(x)} = \frac{{}^nC_{x+1} p^{x+1} q^{n-x-1}}{{}^nC_x\, p^x q^{n-x}}$$

$$= \frac{\underline{|n}\ \underline{|x}\ \underline{|n-x}}{\underline{|x+1}\ \underline{|n-x-1}\underline{|n}} \times \frac{p}{q} \qquad \left[\because\ {}^nC_{x+1} = \frac{\underline{|n}}{\underline{|x+1}\underline{|n-x-1}} \text{ and } {}^nC_x = \frac{\underline{|n}}{\underline{|x}\underline{|n-x}}\right]$$

$$= \frac{\underline{|x}(n-x)\underline{|n-x-1}}{(x+1)\underline{|x}\ \underline{|n-x-1}} \times \frac{p}{q} = \frac{n-x}{x+1} \times \frac{p}{q}$$

$$\Rightarrow p(x+1) = \frac{n-x}{x+1}\frac{p}{q}p(x) \qquad \ldots (3)$$

Putting $x = 0$, 1, 2, 3,…in this equation, we get p(1) in terms of p(0), p(2) in terms of p(1), p(3) in terms of p(2), and so on. Thus, if p(0) is known, we can find p(1) then p(2), p(3) and so on.

So, eqn. (3) is the recurrence relation for finding the probabilities of binomial distribution. The initial probability i.e. p(0) is obtained from the following formula:

$$p(0) = q^n$$

[$\because$ $p(x) = {}^nC_x p^x q^{n-x}$ putting x = 0, we have p(0) = ${}^nC_0 p^0 q^n = q^n$]

Binomial Frequency Distribution

We have studied that in a random experiment with n trials and having p as the probability of success in each trial,

$$P[X = x] = {}^nC_x p^x q^{n-x}; x = 0, 1, 2, ..., n$$

where x is the number of successes. Now, if such a random experiment of n trials is repeated say N times, then the expected (or theoretical) frequency of getting x successes is given by

$f(x) = N.P[X = x] = N.{}^nC_x p^x q^{n-x}; x = 0, 1, 2, ..., n$

i.e. probability is multiplied by N to get the corresponding expected frequency.

Process of Fitting a Binomial Distribution

Suppose we are given the observed frequency distribution. We first find the mean from the given frequency distribution and equate it to np. From this, we can find the value of p. After having obtained the value of p, we obtain $p(0) = q^n$, where q = 1 – p.

Then the recurrence relation i.e. $p(x+1) = \frac{n-x}{x+1} p(x)$ is applied to find the values of p(1), p(2),…. After that, the expected (theoretical) frequencies f(0), (1), f(2), … are obtained on multiplying each of the corresponding probabilities i.e. p(0), p(1), p(2), … by N.

In this way, the binomial distribution is fitted to the given data. Thus, fitting of a binomial distribution involves comparing the observed frequencies with the expected frequencies to see how best the observed results fit with the theoretical (expected) results.

Example 7: Four coins were tossed and number of heads noted. The experiment is repeated 200 times.

The number of tosses showing 0, 1, 2, 3 and 4 heads were found distributed as under. Fit a binomial distribution to these observed results assuming that the nature of the coins is not known.

Number of heads:	0	1	2	3	4
Number of tosses	15	35	90	40	20

Solution: Here n = 4, N = 200.

First, we obtain the mean of the given frequency distribution as follows:

Number of head X	Number of tosses f	fX
0	15	0
1	35	35
2	90	180
3	40	120
4	20	80
Total	200	415

$$\therefore \text{ Mean} = \frac{\sum f(x)}{\sum f} \quad \text{[See Unit 1of MST-002]}$$

$$= \frac{415}{200}$$

$$= 2.075$$

As mean for binomial distribution is np,

$$\therefore \text{ np} = 2.075$$

$$\Rightarrow p = \frac{2.075}{4}$$

$$= 0.5188$$

$$\Rightarrow q = 1 - p$$

$$= 1 - 0.5188$$

$$= 0.4812$$

$$\therefore \text{ p(0)} = q^n$$

$$= (0.4812)^4$$

$$= 0.0536$$

Now, using the recurrence relation

$$p(x+1) = \frac{n-x}{x+1}.\frac{p}{q}p(x); x = 0, 1, 2, 3, 4;$$

we obtain the probabilities for different values of the random variable X i.e. p(1) is obtained on multiplying p(0) with $\frac{4-0}{0+1}$, p(2) is obtained on multiplying p(1) with $\frac{4-1}{1+1}$, and so on; i.e. the values as shown in col. 3 of the following table are obtained on multiplying the preceding values of col. 2 and col 3, except the first value which has been obtained using p(0) = q^n as above.

Number of Heads (X) (1)	$\frac{n-x}{x+1}.\frac{p}{q}=\frac{4-x}{x+1}\left(\frac{0.5188}{0.4812}\right)$ $=\frac{4-x}{x+1}(1.07814)$ (2)	$p(x)$ (3)	Expected or theoretical frequency $f(x)$ (4)
0	$\frac{4-0}{0+1}(1.07814)=4.31256$	p(0) = 0.0536	10.72≈ 11
1	$\frac{4-1}{1+1}(1.07814)=1.61721$	p(1) = 4.31256×0.0536 = 0.23115	46.23≈ 46
2	$\frac{4-2}{2+1}(1.07814)=.071876$	p(2) =1.61721×0.23115 = 0.37382	74.76≈ 75
3	$\frac{4-3}{3+1}(1.07814)=0.26954$	p(3) = 0.71876 × 0.37382 = 0.26869	53.73 ≈ 54
4	$\frac{4-4}{4+1}(1.07814)=0$	p(4) = 0.26954 × .26869 = 0.0724	14.48 ≈ 14

Remark 3: In the above example, if the nature of the coins had been known e.g. if it had been given that "the coins are unbiased" then we would have taken

p = $\frac{1}{2}$ and then the observed data would not have been used to find p. Such a situation can be seen in the problem **E6).**

Here are two exercises for you:

E6) Seven coins are tossed and number of heads noted. The experiment is repeated 128 times and the following distribution is obtained:

Number of heads	0	1	2	3	4	5	6	7
Frequencies	7	6	19	35	30	23	7	1

Fit a binomial distribution assuming the coin is unbiased.

E7) Out of 800 families with 4 children each, how many families would you expect to have 3 boys and 1 girl, assuming equal probability of boys and girls?

Now before ending this unit, let's summarize what we have covered in it.

9.6 SUMMARY

The following main points have been covered in this unit:

1) A discrete random variable X is said to follow **Bernoulli distribution** with parameter p if its probability mass function is given by

$$P[X = x] = \begin{cases} p^x (1-p)^{1-x} & ; x = 0, 1 \\ 0; & \text{elsewhere} \end{cases}$$

Its **mean** and **variance** are **p** and **p(1 − p)**, respectively. **Third** and **fourth central moments** of this distribution are $p(1-p)(1-2p)$ and $p(1-p)(1-3p+3p^2)$ respectively.

2) A discrete random variable X is said to follow **binomial distribution** if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} {}^nC_x p^x q^{n-x} & ; x = 0, 1, 2, ..., n \\ 0; & \text{elsewhere} \end{cases}$$

where, n is the number of independent trials,

x is the number of successes in n trial,

p is the probability of success in each trial, and

q = 1 – p is the probability of failure in each trial.

3) The **constants of Binomial distribution** are:

Mean= np, Variance= npq,

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

$\beta_1 = \frac{(q-p)^2}{npq}$, $\beta_2 = 3 + \frac{1-6pq}{npq}$,

$\gamma_1 = \frac{1-2p}{\sqrt{npq}}$, and $\gamma_2 = \frac{1-6pq}{npq}$

4) For a binomial distribution, **Mean > Variance.**

5) **Recurrence relation for the probabilities of binomial distribution** is

$$p(x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} \cdot p(x), \ x = 0, 1, 2, \ldots, n-1$$

6) The **expected frequencies of the binomial distribution** are given by

$$f(x) = N.P[X = x] = N.\,{}^nC_x p^x q^{n-x}; x = 0, 1, 2, ..., n$$

9.7 SOLUTIONS/ANSWERS

E1) Let p be the probability of hitting the target (success) in a trial.

$\therefore$ n = 5, $p = \frac{1}{4}$, $q = 1 - \frac{1}{4} = \frac{3}{4}$,

and hence by binomial distribution, we have

$$P[X = x] = {}^nC_x p^x q^{n-x} = {}^5C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{5-x}; x = 0, 1, 2, 3, 4, 5.$$

$\therefore$ Required probability = $P[X \geq 2]$

$$= P[X = 2] + P[X = 3] + P[X = 4] + P[X = 5]$$

$$= 1 - \left(P[X = 0] + P[X = 1]\right)$$

$$= 1 - \left[{}^5C_0 \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^{5-0} + {}^5C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^{5-1}\right]$$

$$= 1 - \left[\frac{243}{1024} + \frac{405}{1024}\right] = \frac{376}{1024} = \frac{47}{128}$$

E2) Let p be the probability that the dacoit will be killed (success) by a bullet.

$\therefore$ n = 6, p = 0.6, q = 1 – p = 1 – 0.6 = 0.4, and hence by binomial distribution, we have

$$P[X = x] = {}^nC_x\, p^x\, q^{n-x}; x = 0, 1, 2, ..., n$$

$$= {}^6C_x (0.6)^x (0.4)^{6-x}; x = 0, 1, 2, ..., 6.$$

$\therefore$ The required probability = P[The dacoit is still alive]

= P[No bullet kills the dacoit]

= P[Number of successes is zero]

$= P[X = 0] = {}^6C_0 (0.6)^0 (0.4)^6$

= 0.0041

E3) Mean = np = 3 ... (1)

Variance = npq = 4 ... (2)

$\therefore$ Dividing (2) by (1), we have

$q = \frac{4}{3} > 1$ and hence not possible

[$\because$ q, being probability, cannot be greater than 1]

E4) Let $X \sim B(n, p)$, then

n = 5 and

$np + npq = 4.8$ [$\because$ given that Mean + Variance = 4.8]

$\Rightarrow 5p + 5pq = 4.8$

$\Rightarrow 5[p + p(1-p)] = 4.8$

$\Rightarrow 5[p + p - p^2] = 4.8$

$\Rightarrow 5p^2 - 10p + 4.8 = 0$

$\Rightarrow 25p^2 - 50p + 24 = 0$ [Multiplying by 5]

$\Rightarrow 25p^2 - 30p - 20p + 24 = 0$

$\Rightarrow 5p(5p - 6) - 4(5p - 6) = 0$

$\Rightarrow (5p - 6)(5p - 4) = 0$

$\Rightarrow p = \frac{6}{5}, \frac{4}{5}$

The first value $p = \frac{6}{5}$ is rejected [$\because$ probability can never exceed 1]

$\therefore\ p = \frac{4}{5}$ and hence $q = 1 - p = \frac{1}{5}$.

Thus, the binomial distribution is

$$P[X = x] = {}^nC_x p^x q^{n-x}$$

$$= {}^5C_x \left(\frac{4}{5}\right)^x \left(\frac{1}{5}\right)^{5-x}\ ;\ x = 0, 1, 2, 3, 4, 5.$$

The binomial distribution in tabular form is given as

X	p(x)
0	${}^5C_0 \left(\frac{4}{5}\right)^0 \left(\frac{1}{5}\right)^5 = \frac{1}{3125}$
1	${}^5C_1 \left(\frac{4}{5}\right)^1 \left(\frac{1}{5}\right)^4 = \frac{20}{3125}$
2	${}^5C_2 \left(\frac{4}{5}\right)^2 \left(\frac{1}{5}\right)^3 = \frac{160}{3125}$
3	${}^5C_3 \left(\frac{4}{5}\right)^3 \left(\frac{1}{5}\right)^2 = \frac{640}{3125}$
4	${}^5C_4 \left(\frac{4}{5}\right)^4 \left(\frac{1}{5}\right)^1 = \frac{1280}{3125}$
5	${}^5C_5 \left(\frac{4}{5}\right)^5 \left(\frac{1}{5}\right)^0 = \frac{1024}{3125}$

E5) Given that Mean = 30 and S.D. = 5

Thus, $np = 30$, $\sqrt{npq} = 5$

$\Rightarrow np = 30,\ npq = 25$

i) $\frac{npq}{np}=\frac{25}{30}=\frac{5}{6}\Rightarrow q=\frac{5}{6}, p=1-q=1-\frac{5}{6}=\frac{1}{6},\ n\left(\frac{1}{6}\right)=30\Rightarrow n=180$

ii) $\mu_2=npq=180\times\frac{1}{6}\times\frac{5}{6}=25$

$$\mu_3=npq(q-p)=25\left(\frac{5}{6}-\frac{1}{6}\right)=\frac{50}{3}$$

$$\Rightarrow \beta_1=\frac{\mu_3^2}{\mu_2^3}=\frac{4}{225}$$

$\therefore$ Moment coefficient of skewness is given by

$$\gamma_1=\sqrt{\beta_1}=\frac{2}{15}$$

iii) $\beta_2=3+\frac{1-6pq}{npq}=3+\frac{1-6\times\frac{1}{6}\times\frac{5}{6}}{25}\ =3+\frac{1}{150}$

$$\Rightarrow \gamma_2=\beta_2-3=\frac{1}{150}>0$$

So, the curve of the binomial distribution is leptokurtic.

E6) As the coin is unbiased, $\therefore\ p=\frac{1}{2}$.

Here, n = 7, N = 128, $p=\frac{1}{2}, q=1-p=\frac{1}{2}$.

$\Rightarrow$ p(0) = $q^n=\left(\frac{1}{2}\right)^7=\frac{1}{128}$.

Expected frequencies are, therefore, obtained as follows:

Number of heads (X)	$\frac{n-x}{x+1}.\frac{p}{q}=\frac{7-x}{x+1}.\frac{\frac{1}{2}}{\frac{1}{2}}$ $=\frac{7-x}{x+1}$	$p(x)$	Expected or theoretical Frequency $f(x)=N.p(x)$ $=128.p(x)$
0	$\frac{7-0}{0+1}=7$	$\frac{1}{128}$	1
1	$\frac{7-1}{1+1}=3$	$7\times\frac{1}{128}=\frac{7}{128}$	7
2	$\frac{7-2}{2+1}=\frac{5}{3}$	$3\times\frac{7}{128}=\frac{21}{128}$	21
3	$\frac{7-3}{3+1}=1$	$\frac{5}{3}\times\frac{21}{128}=\frac{35}{128}$	35

4	$\frac{7-4}{4+1}=\frac{3}{5}$	$1\times\frac{35}{128}=\frac{35}{128}$	35
5	$\frac{7-5}{5+1}=\frac{1}{3}$	$\frac{3}{5}\times\frac{35}{128}=\frac{21}{128}$	21
6	$\frac{7-6}{6+1}=\frac{1}{7}$	$\frac{1}{3}\times\frac{21}{128}=\frac{7}{128}$	7
7	$\frac{7-7}{7+1}=0$	$\frac{1}{7}\times\frac{7}{128}=\frac{1}{128}$	1

E7) Here, probability (p) to have a boy is $\frac{1}{2}$ and the probability (q) to have a girl is $\frac{1}{2}$, n = 4, N = 800.

Let X be the number of boys in a family.

$\therefore$ by binomial distribution, the probability of having 3 boys in a family of 4 children

$= P[X = 3] \qquad [\because\ P[X = x] = {}^nC_x p^x q^{n-x}]$

$$= {}^4C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{4-3} = 4\left(\frac{1}{2}\right)^4$$

Hence, the expected number of families having 3 boys and 1 girl

$$= N.p(3) = 128\left(\frac{1}{4}\right) = 32$$

UNIT 10 POISSON DISTRIBUTION

Structure

10.1 INTRODUCTION

In Unit 9, you have studied binomial distribution which is applied in the cases where the probability of success and that of failure do not differ much from each other and the number of trials in a random experiment is finite. However, there may be practical situations where the probability of success is very small, that is, there may be situations where the event occurs rarely and the number of trials may not be known. For instance, the number of accidents occurring at a particular spot on a road everyday is a rare event. For such rare events, we cannot apply the binomial distribution. To these situations, we apply Poisson distribution. The concept of Poisson distribution was developed by a French mathematician, Simeon Denis Poisson (1781-1840) in the year 1837.

In this unit, we define and explain Poisson distribution in Sec. 10.2. Moments of Poisson distribution are described in Sec. 10.3 and the process of fitting a Poisson distribution is explained in Sec. 10.4.

Objectives

After studing this unit, you would be able to:

- know the situations where Poisson distribution is applied;
- define and explain Poisson distribution;
- know the conditions under which binomial distribution tends to Poisson distribution;
- compute the mean, variance and other central moments of Poisson distribution;
- obtain recurrence relation for finding probabilities of this distribution; and
- know as to how a Poisson distribution is fitted to the observed data.

10.2 POISSON DISTRIBUTION

In case of binomial distributions, as discussed in the last unit, we deal with events whose occurrences and non-occurrences are almost equally important. However, there may be events which do not occur as outcomes of a definite number of trials of an experiment but occur rarely at random points of time and for such events our interest lies only in the number of occurrences and not in its non-occurrences. Examples of such events are:

i) Our interest may lie in how many printing mistakes are there on each page of a book but we are not interested in counting the number of words without any printing mistake.

ii) In production where control of quality is the major concern, it often requires counting the number of defects (and not the non-defects) per item.

iii) One may intend to know the number of accidents during a particular time interval.

Under such situations, binomial distribution cannot be applied as the value of n is not definite and the probability of occurrence is very small. Other such situations can be thought of yourself. Poisson distribution discovered by S.D. Poisson (1781-1840) in 1837 can be applied to study these situations.

Poisson distribution is a limiting case of binomial distribution under the following conditions:

i) n, the number of trials is indefinitely large, i.e. $n \to \infty$.

ii) p, the constant probability of success for each trial is very small, i.e. $p \to 0$.

iii) np is a finite quantity say 'λ'.

Definition: A random variable X is said to follow Poisson distribution if it assumes indefinite number of non-negative integer values and its probability mass function is given by:

$$p(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}.\lambda^{x}}{\lfloor x}; & x = 0, 1, 2, 3, \ldots \text{ and } \lambda > 0. \\ 0; & \text{elsewhere} \end{cases}$$

where e = base of natural logarithm, whose value is approximately equal to 2.7183 corrected to four decimal places. Value of $e^{-\lambda}$ can be written from the table given in the Appendix at the end of this unit, or, can be seen from any book of log tables.

Remark 1

i) If X follows Poisson distribution with parameter λ then we shall use the notation $X \sim P(\lambda)$.

ii) If X and Y are two independent Poisson variates with parameters λ_1 and λ_2 repectively, then X + Y is also a Poisson variate with parameter $\lambda_1+\lambda_2$. This is known as **additive property of Poisson distribution**.

10.3 MOMENTS OF POISSON DISTRIBUTION

r^{th} order moment about origin of Poisson variate is

$$\mu_r' = E\left(X^r\right) = \sum_{x=0}^{\infty} x^r p(x) = \sum_{x=0}^{\infty} x^r \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$\mu_1' = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\lfloor x} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x\lfloor x-1} = \sum_{x=1}^{\infty} \frac{e^{-\lambda}.\lambda^x}{\lfloor x-1}$$

$$= e^{-\lambda}\left[\frac{\lambda^1}{\lfloor 0} + \frac{\lambda^2}{\lfloor 1} + \frac{\lambda^3}{\lfloor 2} + \dots\right]$$

$$= \lambda e^{-\lambda}\left[1 + \frac{\lambda^1}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \dots\right]$$

$$= e^{-\lambda}\lambda e^{\lambda} \quad \left[\because e^{\lambda} = 1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^3}{\lfloor 3} + \dots(\text{see Unit 2 of MST-001})\right]$$

$$= \lambda$$

$\therefore$ Mean = λ

$$\mu_2' = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=0}^{\infty}\left[x(x-1)+x\right]\frac{e^{-\lambda}.\lambda^x}{\lfloor x} \qquad \text{[As done in Unit 9 of this Course]}$$

$$= \sum_{x=0}^{\infty}\left(x(x-1)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + x\frac{e^{-\lambda}\lambda^x}{\lfloor x}\right)$$

$$= \sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x(x-1)\lfloor x-2} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^x}{\lfloor x-2} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\left[\frac{\lambda^2}{\lfloor 0} + \frac{\lambda^3}{\lfloor 1} + \frac{\lambda^4}{\lfloor 2} + \dots\right] + \mu_1'$$

$$= e^{-\lambda}\lambda^2\left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \dots\right] + \mu_1'$$

$$= e^{-\lambda}\lambda^2 e^{\lambda} + \mu_1'$$

$$= \lambda^2 + \lambda$$

$\therefore$ Variance of X is given as V(X) = $\mu_2 = \mu_2' - \left(\mu_1'\right)^2$

$$= \lambda^2 + \lambda - (\lambda)^2$$

$$= \lambda$$

$$\mu_3' = \sum_{x=0}^{3} x^3 p(x)$$

Writing x^3 as $x(x-1)(x-2)+3x(x-1)+x$, we have

$$\begin{bmatrix}\text{See Unit 9 of this course where} \\ \text{the expression of } \mu_3' \text{ is obtained}\end{bmatrix}$$

$$= \sum_{x=0}^{\infty}\left[x(x-1)(x-2)+3x(x-1)+x\right]\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=3}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + 3\sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + \sum_{x=1}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\sum_{x=3}^{\infty} x(x-1)(x-2)\frac{\lambda^x}{x(x-1)(x-2)\lfloor x-3} + 3(\lambda^2) + (\lambda)$$

$$= e^{-\lambda}\sum_{x=3}^{\infty}\frac{\lambda^x}{\lfloor x-3} + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\left(\frac{\lambda^3}{\lfloor 0} + \frac{\lambda^4}{\lfloor 1} + \frac{\lambda^5}{\lfloor 2} + \ldots\right) + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\lambda^3\left(1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \ldots\right) + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3\lambda^2 + \lambda$$

$$= \lambda^3 + 3\lambda^2 + \lambda$$

Third order central moment is

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

$$= \lambda \qquad \text{[On simplification]}$$

$$\mu_4' = \sum_{x=3}^{\infty} x^4 . \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

Now writing $x^4 = x(x-1)(x-2)(x-3)+6x(x-1)(x-2)+7x(x-1)+x$, and proceeding in the similar fashion as done in case of μ_3', we have

$$\mu_4' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

$\therefore$ Fourth order central moment is

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

$$= 3\lambda^2 + \lambda \qquad \text{[On simplification]}$$

Therefore, measures of skewness and kurtosis are given by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}, \qquad \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}; \text{ and}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2 + \lambda}{\lambda^2} = 3 + \frac{1}{\lambda}, \qquad \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}.$$

Now as γ_1 is positive, therefore the Poisson distribution is always positively skewed distribution. Also as γ_2 > ($\because$ $\lambda > 0$), the curve of the distribution is Leptokurtic.

Remark 2

i) Mean and variance of Poisson distribution are always equal. In fact this is the only discrete distribution for which Mean = Variance = the third central moment.

ii) Moments of the Poisson distribution can be deduced from those of the binomial distribution also as explained below:

For a binomial distribution,

Mean = np

Variance = npq

$\mu_3 = \text{npq}(q - p)$

$\mu_4 = \text{npq}\left[1 + 3\text{pq}(n-2)\right] = \text{npq}\left[1 + 3\text{npq} - 6\text{pq}\right]$

Now as the Poisson distribution is a limiting form of binomial distribution under the conditions:

(i) $n \to \infty$, (ii) $p \to 0$ i.e. $q \to 1$, and (iii) $np = \lambda$ (a finite quantity);

$\therefore$ Mean, Variance and other moments of the Poisson distribution are given as:

Mean = Limiting value of np = λ

Variance = Limiting value of npq

= Limiting value of (np) (q)

= (λ) (1) = λ

μ_3 = Limiting value of npq (q – p)

= Limiting value of (npq) (q – p)

= (λ) (1 – 0)

= λ

μ_4 = Limiting value of npq [1 + 3npq – 6pq]

= Limiting value of (npq) [1 + 3(npq) –6 (p)(q)]

= (λ) [1 + 3(λ) – 6 (0)(1)]

= $\lambda[1 + 3\lambda] = 3\lambda^2 + \lambda$

Now let’s give some examples of Poisson distribution.

Example 1: It is known that the number of heavy trucks arriving at a railway station follows the Poisson distribution. If the average number of truck arrivals during a specified period of an hour is 2, find the probabilities that during a given hour

a) no heavy truck arrive,

b) at least two trucks will arrive.

Solution: Here, the average number of truck arrivals is 2

i.e. mean = 2

$\Rightarrow \lambda = 2$

Let X be the number of trucks arrive during a given hour,

$\therefore$ by Poisson distribution, we have

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{\lfloor x} = \frac{e^{-2}(2)^x}{\lfloor x}; x = 0, 1, 2, ...$$

Thus, the desired probabilities are:

(a) P[arrival of no heavy truck] = P[X = 0]

$$= \frac{e^{-2}2^0}{\lfloor 0}$$

$$= e^{-2}$$

$$= 0.1353 \quad \begin{bmatrix}\text{See the table given} \\ \text{in the Appendix at} \\ \text{the end of this unit}\end{bmatrix}$$

(b) P[arrival of at least two trucks] = $P[X \geq 2]$

$$= P[X = 2] + P[X = 3] + ...$$

$$= 1 - \left[P[X = 1] + P[X = 0]\right]$$

$$\begin{bmatrix}\because \text{ sum of all the} \\ \text{probabilities is 1}\end{bmatrix}$$

$$= 1 - \left[\frac{e^{-2}2^0}{\lfloor 0} + \frac{e^{-2}2^1}{\lfloor 1}\right]$$

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

$$= 1 - (0.1353)(3) = 1 - 0.4059 = 0.5941$$

Note: In most of the cases for Poisson distribution, if we are to compute the probabilities of the type $P[X > a]$ or $P[X \geq a]$, we write them as

$P[X > a] = 1 - P[X \leq a]$ and

$P[X \geq a] = 1 - P[X < a]$, because n may not be definite and hence we cannot go up to the last value and hence the probability is written in terms of its complementary probability.

Example 2: If the probability that an individual suffers a bad reaction from an injection of a given serum is 0.001, determine the probability that out of 500 individuals

i) exactly 3,

ii) more than 2

individuals suffer from bad reaction

Solution: Let X be the Poisson variate, "Number of individuals suffering from bad reaction". Then,

n = 1500, p = 0.001,

$\therefore$ λ = np = (1500) (0.001) = 1.5

$\therefore$ By Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda}\lambda^{x}}{\underline{|x}}, x = 0, 1, 2, ...$$

$$= \frac{e^{-1.5}.(1.5)^{x}}{\underline{|x}}; x = 0, 1, 2, ...$$

Thus,

i) The desired probability = P[X = 3]

$$= \frac{e^{-1.5}.(1.5)^{3}}{\underline{|3}}$$

$$= \frac{(0.2231)(3.375)}{6} \quad = 0.1255$$

$$\begin{bmatrix} \because e^{-0.5} = 0.6065, e^{-1} = 0.3679, \text{so} \\ e^{-1.5} = e^{-1} \times e^{-0.5} = (0.3679)(0.6065) = 0.2231) \\ \text{See the table given in the Appendix} \\ \text{at the end of this unit} \end{bmatrix}$$

ii) The desired probability $= P[X > 2]$

$$= 1 - P[X \leq 2]$$

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

$$= 1 - \left[\frac{e^{-1.5}.(1.5)^{2}}{\underline{|2}} + \frac{e^{-1.5}.(1.5)^{1}}{\underline{|1}} + \frac{e^{-1.5}.(1.5)^{0}}{\underline{|0}}\right]$$

$$= 1 - e^{-1.5}\left[\frac{2.25}{2} + 1.5 + 1\right] = 1 - (3.625)e^{-1.5}$$

$$= 1 - (3.625)(0.2231) = 1 - 0.8087 = 0.1913$$

Example 3: If the mean of a Poisson distribution is 1.44, find the values of variance and the central moments of order 3 and 4.

Solution: Here, mean = 1.44

$\Rightarrow \lambda = 1.44$

Hence, Variance = λ = 1.44

$\mu_3 = \lambda = 1.44$

$\mu_4 = 3\lambda^2 + \lambda = 3\,(1.44)^2 + 1.44 = 7.66.$

Example 4: If a Poisson variate X is such that P[X = 1] = 2P[X = 2], find the mean and variance of the distribution.

Solution: Let λ be the mean of the distribution, hence by Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{\lfloor x}; x = 0, 1, 2, \ldots$$

Now, $P[X = 1] = 2P[X = 2]$

$$\Rightarrow \frac{e^{-\lambda}.\lambda^1}{\lfloor 1} = 2\frac{e^{-\lambda}.\lambda^2}{\lfloor 2}$$

$\Rightarrow \lambda = \lambda^2 \Rightarrow \lambda^2 - \lambda = 0 \Rightarrow \lambda(\lambda - 1) = 0 \Rightarrow \lambda = 0, 1$

But $\lambda = 0$ is rejected

[$\because$ if $\lambda = 0$ then either n = 0 or p = 0 which implies that Poisson distribution does not exist in this case.]

$\therefore \lambda = 1$

Hence mean = λ = 1, and

Variance = λ = 1.

Example 5: If X and Y be two independent Poisson variates having means 1 and 2 respectively, find P[X + Y < 2].

Solution: As X ~ P(1), Y ~ P(2), therefore,

X + Y follows Poisson distribution with mean = 1 + 2 = 3.

Let X + Y = W. Hence, probability function of W is

$$P[W = w] = \frac{e^{-3}.3^w}{\lfloor w}; w = 0, 1, 2, \ldots.$$

Thus, the required probability= P[X + Y < 2]

= P[W < 2]

= P[W = 0] + P[W = 1]

$$= \frac{e^{-3}.3^0}{\lfloor 0} + \frac{e^{-3}.3^1}{\lfloor 1}$$

$= (0.0498)(1 + 3)$ [From Table, $e^{-3} = 0.0498$]

$= 0.1992.$

You may now try these exercises.

E1) Assume that the chance of an individual coal miner being killed in a mine accident during a year is $\frac{1}{1400}$. Use the Poisson distribution to calculate the probability that in a mine employing 350 miners, there will be at least one fatal accident in a year. (use $e^{-0.25} = 0.78$)

E2) The mean and standard deviation of a Poisson distribution are 6 and 2 respectively. Test the validity of this statement.

E3) For a Poisson distribution, it is given that P[X = 1] = P[X = 2], find the value of mean of distribution. Hence find P[X = 0] and P[X = 4].

We now explain as to how the Poisson distribution is fitted to the observed data.

10.4 FITTING OF POISSON DISTRIBUTION

To fit a Poisson distribution to the observed data, we find the theoretical (or expected) frequencies corresponding to each value of the Poisson variate. Process of finding the probabilities corresponding to each value of the Poisson variate becomes easy if we use the recurrence relation for the probabilities of Poisson distribution. So, in this section, we will first establish the recurrence relation for probabilities and then define the Poisson frequency distribution followed by the process of fitting a Poisson distribution.

Recurrence Formula for the Probabilities of Poisson Distribution

For a Poisson distribution with parameter λ, we have

$$p(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x} \quad \ldots (1)$$

Changing x to x + 1, we have

$$p(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{\lfloor x+1} \quad \ldots (2)$$

Dividing (2) by (1), we have

$$\frac{p(x+1)}{p(x)} = \frac{\left(\frac{e^{-\lambda}\lambda^{x+1}}{\lfloor x+1}\right)}{\left(\frac{e^{-\lambda}\lambda^{x}}{\lfloor x}\right)} = \frac{\lambda}{x+1}$$

$$\Rightarrow p(x+1) = \frac{\lambda}{x+1} p(x) \qquad \ldots (3)$$

This is the recurrence relation for probabilities of Poisson distribution. After obtaining the value of p(0) using Poisson probability function i.e.

$p(0) = \frac{e^{-\lambda}\lambda^0}{(0)} = e^{-\lambda}$, we can obtain p(1), p(2), p(3),…, on putting x = 0, 1, 2, …. successively in (3).

Poisson Frequency Distribution

If an experiment, satisfying the requirements of Poisson distribution, is repeated N times, then the expected frequency of getting x successes is given by

$$f(x) = N.P[X = x] = N.\frac{e^{-\lambda}\lambda^x}{\lfloor x};\ x = 0, 1, 2, \ldots$$

Example 5: A manufacturer, who produces medicine bottles, finds that 0.1% 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 at least two defective bottles.

Solution: Let X be the Poisson variate, "the number of defective bottles in a box". Here, number of bottles in a box (n) = 500, therefore, the probability (p) of a bottle being defective is

$$p = 0.1\% = \frac{0.1}{100} = 0.001$$

Number of boxes (N) = 100

$$\lambda = np = 500 \times .001 = 0.5$$

Using Poisson distribution, we have

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{\lfloor x};\ x = 0, 1, 2, \ldots$$

$$= \frac{e^{-0.5}(0.5)^x}{\lfloor x};\ x = 0, 1, 2, \ldots$$

$\therefore$ Probability that a box contain at least two defective bottles

$$= P[X \ge 2]$$

$$= 1 - P[X < 2]$$

$$= 1 - \left[P[X = 0] + P[X = 1]\right]$$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{\lfloor 0} + \frac{e^{-0.5}(0.5)^1}{\lfloor 1}\right] = 1 - e^{-0.5}[1 + 0.5]$$

= 1 – (0.6065) (1.5) = 1 – 0.90975 = 0.09025.

Hence, the expected number of boxes containing at least two defective bottles

$= N.P[X \geq 2]$

$= (100)\ (0.09025)$

$= 9.025$

Process of Fitting a Poison Distribution

For fitting a Poisson distribution to the observed data, you are to proceed as described in the following steps.

- First we obtain mean of the given distribution i.e. $\frac{\sum fx}{\sum f}$, being mean, take this as the value of λ.
- Next we obtain $p(0) = e^{-\lambda}$ [Use table given in Appendix at the end of this unit.]
- The recurrence relation $p(x+1) = \frac{\lambda}{x+1} p(x)$ is then used to compute the values of p(1), p(2), p(3), …
- The probabilities obtained in the preceding two steps are then multiplied with N to get expected/theoretical frequencies i.e. $f(x) = N.P[X = x]; x = 0, 1, 2, ...$

Example 6: The following data give frequencies of aircraft accidents experienced by 2480 pilots during a certain period:

Number of Accidents	0	1	2	3	4	5
Frequencies	1970	422	71	13	3	1

Fit a Poisson distribution and calculate the theoretical frequencies.

Solution: Let X be the number of accidents of the pilots. Let us first obtain the mean number of accidents as follows:

Number of Accidents (X)	Frequency (f)	f X
0	1970	0
1	422	422
2	71	142
3	13	39
4	3	12
5	1	5
Total	2480	620

$$\therefore \text{ Mean} = \lambda = \frac{\sum fx}{\sum f} = \frac{620}{2480}$$

$$\Rightarrow \lambda = 0.25$$

$\therefore$ by Poisson distribution,

$$p(0) = e^{-\lambda} = e^{-0.25}$$

$$= 0.7788 \quad \left[\begin{array}{l}\text{See table given in the Appendix} \\ \text{at the end of this unit}\end{array}\right]$$

Now, using the recurrence relation for probabilities of Poisson distribution i.e. $p(x+1) = \frac{\lambda}{x+1} p(x)$ and then multiplying each probability with N, we get the expected frequencies as shown in the following table

Number of Accidents (X) (1)	$\frac{\lambda}{x+1} = \frac{0.25}{x+1}$ (2)	$p(x) = P[X = x]$ (3)	Expected/ Theoretical frequency f(x) = 2480p(x) (4)
0	$\frac{0.25}{0+1} = 0.25$	p(0) = 0.7788	1931. 4 $\approx$ 1931
1	$\frac{0.25}{1+1} = 0.125$	p(1) = 0.25 × 0.7788 = 0.1947	482. 9 $\approx$ 483
2	$\frac{0.25}{2+1} = 0.0833$	p(2) = 0.125 × 0.1947 = 0.0243	60.3 $\approx$ 60
3	$\frac{0.25}{3+1} = 0.0625$	p(3)= 0.0833 × 0.0243 = 0.0020	4.96 $\approx$ 5
4	$\frac{0.25}{4+1} = 0.05$	p(4)= 0.0625 × 0.0020 = 0.0001	0.248 $\approx$ 0
5	$\frac{0.25}{5+1} = 0.0417$	p(5)= 0.05 × 0.0001 = 0.000005	0

You can now try the following exercises

E4) In a certain factory turning out fountain pens, there is a small chance, $\frac{1}{500}$, for any pen to be defective. The pens are supplied in packets of 10. Calculate the approximate number of packets containing (i) one defective (ii) two defective pens in a consignment of 20000 packets.

E5) A typist commits the following mistakes per page in typing 100 pages. Fit a Poisson distribution and calculate the theoretical frequencies.

Mistakes per page(X)	0	1	2	3	4	5
Frequency (f)	42	33	14	6	4	1

We now conclude this unit by giving a summary of what we have covered in it.

10.5 SUMMARY

The following main points have been covered in this unit:

1. A random variable X is said to follow **Poisson distribution** if it assumes indefinite number of non-negative integer values and its probability mass function is given by:

$$p(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{\underline{|x}}; & x = 0, 1, 2, 3, \ldots \text{ and } \lambda > 0. \\ 0; & \text{elsewhere} \end{cases}$$

2. For Poisson distribution, **Mean = Variance = μ_3** = λ, **μ_4**$= 3\lambda^2 + \lambda$

3. $\beta_1 = \dfrac{1}{\lambda}$, $\gamma_1 = \dfrac{1}{\sqrt{\lambda}}$, $\beta_2 = 3 + \dfrac{1}{\lambda}$, $\gamma_2 = \dfrac{1}{\lambda}$ for this distribution .

4. **Recurrence relation for probabilities of Poisson distribution** is

$$p(x+1) = \frac{\lambda}{x+1}.p(x),\ x = 0, 1, 2, 3, \ldots$$

5. **Expected frequencies for a Poisson distribution** are given by

$$f(x) = N.P[X = x] = N.\frac{e^{-\lambda}\lambda^x}{\underline{|x}}; x = 0, 1, 2, \ldots$$

If you want to see what our solutions/answers to the exercises in the unit are, we have given them in the following section.

10.6 SOLUTIONS/ANSWERS

E1) Let X be the Poisson variable "Number of fatal accidents in a year".

Here n = 350, $p = \dfrac{1}{1400}$

$$\Rightarrow \lambda = np = (350)\left(\frac{1}{1400}\right) = 0.25.$$

By Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda}.\lambda^{x}}{\underline{|x}}, x = 0, 1, 2, ...$$

$$= \frac{e^{-0.25}(0.25)^{x}}{\underline{|x}}, x = 0, 1, 2, ...$$

Therefore, P [at least one fatal accident]

$$= P[X \geq 1] = 1 - P[X < 1] = 1 - P[X = 0]$$

$$= 1 - \frac{e^{-0.25}(0.25)^{0}}{\underline{|0}} = 1 - e^{-0.25} = 1 - 0.78 = 0.22$$

E2) As mean = 6, therefore, $\lambda = 6$.

As standard deviation is 2, therefore, variance = 4 $\Rightarrow$ $\lambda = 4$.

We get two different values of λ, which is impossible. Hence, the statement is invalid.

E3) Let λ be the mean of the distribution,

$\therefore$ by Poisson distribution, we have

$$P[X = x] = \frac{e^{-\lambda}\lambda^{x}}{\underline{|x}}; x = 0, 1, 2, 3, ...$$

Given that P[X = 1] = P[X = 2],

$$\therefore \frac{e^{-\lambda}\lambda^{1}}{\underline{|1}} = \frac{e^{-\lambda}\lambda^{2}}{\underline{|2}}$$

$$\Rightarrow \lambda = \frac{\lambda^{2}}{2} \Rightarrow \lambda^{2} - 2\lambda = 0 \Rightarrow \lambda(\lambda - 2) = 0$$

$\Rightarrow \lambda = 0, 2.$

$\lambda = 0$ is rejected,

$\therefore \lambda = 2$

Hence, Mean = 2.

Now, $P[X = 0] = \dfrac{e^{-\lambda}\lambda^{0}}{\underline{|0}} = e^{-\lambda} = e^{-2} = 0.1353,$

[See table given in the Appendix at the end of this unit.]

and $P[X = 4] = \dfrac{e^{-\lambda}\lambda^{4}}{\underline{|4}} = \dfrac{e^{-2}(2)^{4}}{\underline{|4}} = \dfrac{e^{-2}(16)}{24} = \dfrac{2}{3}(0.1353)$

$= 2(0.0451)$

$= 0.0902.$

E4) Here $p = \frac{1}{500},\ n = 10,\ N = 20000,$

$\therefore\ \lambda = np = 10 \times \frac{1}{500} = 0.02$

By Poisson frequency distribution

$f\left(x\right) = N.P\left[X = x\right]$

$= \left(20000\right)\frac{e^{-\lambda}\lambda^{x}}{\underline{|x}};\ x = 0,\ 1,\ 2,...$

Now,

i) The number of packets containing one defective

$= f(1)$

$= \left(20000\right)\frac{e^{-0.02}.\left(0.02\right)^{1}}{\underline{|1}}$

$= (20000)\ (0.9802)\ (0.02)$ $\left[\begin{array}{l}\text{See the table given} \\ \text{in the Appendix}\end{array}\right]$

$= 392.08 \sqcup 392$; and

ii) The number of packets containing two defectives

$= f(2) = 20000\frac{e^{-0.02}\left(0.02\right)^{2}}{\underline{|2}}$

$= \left(20000\right)\frac{\left(0.9802\right)\left(0.0004\right)}{2} = 3.9208 \sqcup 4$

E5) The mean of the given distribution is computed as follows

X	**f**	**fX**
0	42	0
1	33	33
2	14	28
3	6	18
4	4	16
5	1	5
Total	100	100

$\therefore$ Mean $\lambda = \frac{\sum fx}{\sum f} = \frac{100}{100} = 1$

$\Rightarrow\ p\left(0\right) = e^{-\lambda} = e^{-1} = 0.3679.$

Now, we obtain p(1), p(2), p(3), p(4), p(5) using the recurrence relation for probabilities of Poisson distribution i.e.

$p(x+1) = \frac{\lambda}{x+1} p(x); x = 0, 1, 2, 3, 4$ and then obtain the expected frequencies as shown in the following table:

X	$\frac{\lambda}{x+1} = \frac{1}{x+1}$	$p(x)$	Expected/Theoretical frequency $f(x) = N.P(X = x)$ $= 100.P(X = x)$
0	$\frac{1}{0+1} = 1$	$p(0) = 0.3679$	36.79 □ 37
1	$\frac{1}{1+1} = 0.5$	$p(1) = 1 \times 0.3679 = 0.3679$	36.79 □ 37
2	$\frac{1}{2+1} = 0.3333$	$p(2) = 0.5 \times 0.3679 = 0.184$	18.4 □ 18
3	$\frac{1}{3+1} = 0.25$	p(3) = $0.3333 \times 0.184 = 0.0613$	6.13 □ 6
4	$\frac{1}{4+1} = 0.2$	p(4)=$0.25 \times 0.0613 = 0.0153$	1.53 □ 2
5	$\frac{1}{5+1} = 0.1667$	p(5)=$0.2 \times 0.0153 = 0.0031$	0.3 □ 0

Appendix

Value of $e^{-\lambda}$ (For Computing Poisson Probabilities)

(0 < λ < I)

λ	0	1	2	3	4	5	6	7	8	9
0.0	1.0000	0.9900	0.9802	0.9704	0.9608	0.9512	0.9418	0.9324	0.9231	0.9139
0.1	0.9048	0.8958	0.8860	0.8781	0.8694	0.8607	0.8521	0.8437	0.8353	0.8270
0.2	0.7187	0.8106	0.8025	0.7945	0.7866	0.7788	0.7711	0.7634	0.7558	0.7483
0.3	0.7408	0.7334	0.7261	0.7189	0.7118	0.7047	0.6970	0.6907	0.6839	0.6771
0.4	06703	0.6636	0.6570	0.6505	0.6440	0.6376	0.6313	0.6250	0.6188	0.6125
0.5	0.6065	0.6005	0.5945	0.5886	0.5827	0.5770	0.5712	0.5655	0.5599	0.5543
0.6	0.5448	0.5434	0.5379	0.5326	0.5278	0.5220	0.5160	0.5113	0.5066	0.5016
0.7	0.4966	0.4916	0.4868	0.4810	0.4771	0.4724	0.4670	0.4630	0.4584	0.4538
0.8	0.4493	0.4449	0.4404	0.4360	0.4317	0.4274	0.4232	0.4190	0.4148	0.4107
0.9	0.4066	0.4026	0.3985	0.3946	0.3906	0.3867	0.3829	0.3791	0.3753	0.3716
(λ=1, 2, 3, ...,10)										
λ	1	2	3	4	5	6	7	8	9	10
$e^{-\lambda}$	0.3679	0.1353	0.0498	0.0183	0.0070	0.0028	0.0009	0.0004	0.0001	0.00004

Note: To obtain values of $e^{-\lambda}$ for other values of λ, use the laws of exponents i.e.

$e^{-(a+b)} = e^{-a}.e^{-b}$ e. g. $e^{-2.25} = e^{-2}.e^{-0.25} = (0.1353)(0.7788) = 0.1054.$

UNIT 11 DISCRETE UNIFORM AND HYPERGEOMETRIC DISTRIBUTIONS

Structure

11.1 INTRODUCTION

In the previous two units, we have discussed binomial distribution and its limiting form i.e. Poisson distribution. Continuing the study of discrete distributions, in the present unit, two more discrete distributions – Discrete uniform and Hypergeometric distributions are discussed.

Discrete uniform distribution is applicable to those experiments where the different values of random variable are equally likely. If the population is finite and the sampling is done without replacement i.e. if the events are random but not independent, then we use Hypergemetric distribution.

In this unit, discrete uniform distribution and hypergeometric distribution are discussed in Secs. 11.2 and 11.3, respectively. We shall be discussing their properties and applications also in these sections.

Objectives

After studing this unit, you should be able to:

- define the discrete uniform and hypergeometric distributions;
- compute their means and variances;
- compute probabilities of events associated with these distributions; and
- know the situations where these distributions are applicable.

11.2 DISCRETE UNIFORM DISTRIBUTION

Discrete uniform distribution can be conceived in practice if under the given experimental conditions, the different values of the random variable are equally likely. For example, the number on an unbiased die when thrown may be 1 or 2 or 3 or 4 or 5 or 6. These values of random variable, “the number on an unbiased die when thrown” are equally likely and for such an experiment, the discrete uniform distribution is appropriate.

Definition: A random variable X is said to have a discrete uniform (rectangular) distribution if it takes any positive integer value from 1 to n, and its probability mass function is given by

$$P[X = x] = \begin{cases} \frac{1}{n} & \text{for } x = 1, 2, ..., n \\ 0, & \text{otherwise.} \end{cases}$$

where n is called the parameter of the distribution.

For example, the random variable X, "the number on the unbiased die when thrown", takes on the positive integer values from 1 to 6 follows discrete uniform distribution having the probability mass function.

$$P[X = x] = \begin{cases} \frac{1}{6}\ , & \text{for } x = 1, 2, 3, 4, 5, 6. \\ 0\ , & \text{otherwise.} \end{cases}$$

Mean and Variance of the Distribution

$$\text{Mean} = E(X) = \sum_{x=1}^{n} x\, p(x) = \sum_{x=1}^{n} x.\left(\frac{1}{n}\right) = \frac{1}{n}\sum_{x=1}^{n} x$$

$$= \frac{1}{n}[1+2+3+...+n]$$

$$= \frac{1}{n}.\frac{n(n+1)}{2} \quad \begin{bmatrix} \because \text{sum of first n natural numbers} = \frac{n(n+1)}{2} \\ \text{(see Unit 3 of Course MST} - 001) \end{bmatrix}$$

$$= \frac{n+1}{2}.$$

$$\text{Variance} = E(X^2) - [E(X)]^2 \qquad [\because \mu_2 = \mu_2' - (\mu_1')^2]$$

where

$$E(X) = \frac{n+1}{2} \qquad \text{[Obtained above]}$$

$$E(X^2) = \sum_{x=1}^{n} x^2.p(x)$$

and $E(X^2) = \sum_{x=1}^{n} x^2.\frac{1}{n}$

$$= \frac{1}{n}[1^2 + 2^2 + 3^2 + ... + n^2]$$

$$= \frac{1}{n}\left[\frac{n(n+1)(2n+1)}{6}\right] \quad \begin{bmatrix} \because \text{sum of squares of first n} \\ \text{natural numbers} = \frac{n(n+1)(2n+1)}{6} \\ \text{(see Unit 3 of Course MST} - 001) \end{bmatrix}$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$\therefore \text{ Variance} = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$

$$= \frac{(n+1)}{12}\left[2(2n+1) - 3(n+1)\right]$$

$$= \frac{n+1}{12}[4n+2-3n-3] = \frac{(n+1)}{12}(n-1) = \frac{n^2-1}{12}$$

Example 1: Find the mean and variance of a number on an unbiased die when thrown.

Solution: Let X be the number on an unbiased die when thrown,

$\therefore$ X can take the values 1, 2, 3, 4, 5, 6 with

$$P[X = x] = \frac{1}{6}; x = 1, 2, 3, 4, 5, 6.$$

Hence, by uniform distribution, we have

$$\text{Mean} = \frac{n+1}{2} = \frac{6+1}{2} = \frac{7}{2}, \text{ and}$$

$$\text{Variance} = \frac{n^2-1}{12} = \frac{(6)^2-1}{12} = \frac{35}{12}.$$

Uniform Frequency Distribution

If an experiment, satisfying the requirements of discrete uniform distribution, is repeated N times, then expected frequency of a value of random variable is given by

$$f(x) = N.P[X = x]; x = 1, 2, ..., n$$

$$= N.\frac{1}{n}; x = 1, 2, 3, ..., n.$$

Example 2: If an unbiased die is thrown 120 times, find the expected frequency of appearing 1, 2, 3, 4, 5, 6 on the die.

Solution: Let X be the uniform discrete random variable, "the number on the unbiased die when thrown".

$$\therefore P[X = x] = \frac{1}{6}; x = 1, 2, ..., 6$$

Hence, the expected frequencies of the value of random variable are given as computed in the following table:

X	$P[X=x]$	Expected/Theoretical frequencies $f(x) = N.P[X=x] = 120.P[X=x]$
1	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
2	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
3	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
4	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
5	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
6	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$

Now, you can try the following exercise:

E1) Obtain the mean, variance of the discrete uniform distribution for the random variable, "the number on a ticket drawn randomly from an urn containing 10 tickets numbered from 1 to 10". Also obtain the expected frequencies if the experiment is repeated 150 times.

11.3 HYPERGEOMETRIC DISTRIBUTION

In the last section of this unit, we have studied discrete uniform probability distribution wherein the probability distribution is obtained for the possible outcomes in a single trial like drawing a ticket from an urn containing 10 tickets as mentioned in exercise **E1).** But, if there are more than one but finite trials with only two possible outcomes in each trial, we apply some other distribution. One such distribution which is applicable in such a situation is binomial distribution which you have studied in Unit 9. The binomial distribution deals with finite and independent trials, each of which has exactly two possible outcomes (Success or Failure) with constant probability of success in each trial. For example, if we again consider the example of drawing ticket randomly from an urn containing 10 tickets bearing numbers from 1 to 10. Then, the probability that the drawn ticket bears an odd number is $\frac{5}{10} = \frac{1}{2}$. If we replace the ticket back, then the probability of drawing a ticket bearing an odd number is again $\frac{5}{10} = \frac{1}{2}$. So, if we draw ticket again and again with replacement, trials become independent and probability of getting an odd number is same in each trial. Suppose, it is asked that what is the probability of getting 2 tickets bearing odd number in 3 draws then we apply binomial distribution as follows:

Let X be the number of times on odd number appears in 3 draws, then by binomial distribution,

$$P[X=2] = {}^3C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{3-2} = (3)\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \frac{3}{8}.$$

But, if in the example discussed above, we do not replace the ticket after any draw the probability of getting an odd number gets changed in each trial and the trials remain no more independent and hence in this case binomial distribution is not applicable. Suppose, in this case also, we are interested in finding the probability of getting ticket bearing odd number twice in 3 draws, then it is computed as follows:

Let A_i be the event that i^{th} ticket drawn bears odd number and $\bar{A}_i$ be the event that i^{th} ticket drawn does not bear odd number.

$\therefore$ Probability of getting ticket bearing odd number twice in 3 draws

$$= P\left[A_1 \cap A_2 \cap \bar{A}_3\right] + P\left[A_1 \cap \bar{A}_2 \cap A_3\right] + P\left[\bar{A}_1 \cap A_2 \cap A_3\right]$$

[As done in Unit 3 of this Course]

$$= P[A_1]P[A_2 \mid A_1]P\left[\bar{A}_3 \mid A_1 \cap A_2\right] + P[A_1]P\left[\bar{A}_2 \mid A_1\right]P\left[A_3 \mid A_1 \cap \bar{A}_2\right]$$
$$+ P\left[\bar{A}_1\right]P\left[A_2 \mid \bar{A}_1\right]P\left[A_3 \mid \bar{A}_1 \cap A_2\right]$$

[Multiplication theorem for dependent events (See Unit 3 of this Course)]

$$= \frac{5}{10}.\frac{4}{9}.\frac{5}{8} + \frac{5}{10}.\frac{5}{9}.\frac{4}{8} + \frac{5}{10}.\frac{5}{9}.\frac{4}{8}$$

$$= 3 \times \frac{5\times5\times4}{10\times9\times8}$$

This result can be written in the following form also:

$$= \frac{5\times4\times5\times3\times2}{2\times10\times9\times8} \qquad \text{[Multiplying and Dividing by 2]}$$

$$= \frac{5\times4}{2}\times5\times\frac{1}{\frac{10\times9\times8}{3\times2}} = {}^5C_2\times{}^5C_1\times\frac{1}{{}^{10}C_3} = \frac{{}^5C_2\times{}^5C_1}{{}^{10}C_3}$$

In the above result, ${}^5C_2$ is representing the number of ways of selecting 2 out of 5 tickets bearing odd number, ${}^5C_1$ is representing the number of ways of selecting 1 out of 5 tickets bearing even number i.e. not bearing odd number, and ${}^{10}C_3$ is representing the number of ways of selecting 3 out of total 10 tickets.

Let us consider another similar example of a bag containing 20 balls out of which 5 are white and 15 are black. Suppose 10 balls are drawn at random one by one without replacement, then as discussed in the above example, the probability that in these 10 draws, there are 2 white and 8 black balls is

$$\frac{{}^{5}C_2 \times {}^{15}C_8}{{}^{20}C_{10}}.$$

Note: The result remains exactly same whether the items are drawn one by one without replacement or drawn at once.

Let us now generalize the above argument for N balls, of which M are white and $N - M$ are black. Of these, n balls are chosen at random without replacement. Let X be a random variable that denote the number of white balls drawn. Then, the probability of $X = x$ white balls among the n balls drawn is given by

$$P[X = x] = \frac{{}^{M}C_x . {}^{N-M}C_{n-x}}{{}^{N}C_n}$$

[For $x = 0, 1, 2, ..., n\,(n \le M)$ or $x = 0, 1, 2, ..., M\,(n > M)$]

The above probability function of discrete random variable X is called the Hypergeometric distribution.

Remark 1: We have a hypergeometric distribution under the following conditions:

i) There are finite number of dependent trials

ii) A single trial results in one of the two possible outcomes-Success or Failure

iii) Probability of success and hence that of failure is not same in each trial i.e. sampling is done without replacement

Remark 2: If number (n) of balls drawn is greater than the number (M) of white balls in the bag, then if $n \le M$, the number (x) of white balls drawn cannot be greater than n and if $n > M$, then number of white balls drawn cannot be greater than M. So, x can take the values upto n $(\text{if } n \le M)$ and M(if $n > M$) i.e. x can take the value upto n or M, whichever is less, i.e. $x = \min\{n, M\}$.

The discussion leads to the following definition

Definition: A random variable X is said to follow the hypergeometric distribution with parameters N, M and n if it assumes only non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} \dfrac{{}^{M}C_x . {}^{N-M}C_{n-x}}{{}^{N}C_n} & \text{for } x = 0, 1, 2, ..., \min\{n, M\} \\ 0, & \text{otherwise} \end{cases}$$

where n, M, N are positive integers such that $n \le N$, $M \le N$.

Mean and Variance

$$\text{Mean} = E(X) = \sum_{x=0}^{n} x.p[X = x]$$

$$= \sum_{x=1}^{n} x.\frac{{}^{M}C_x . {}^{N-M}C_{n-x}}{{}^{N}C_n}$$

$$= \sum_{x=1}^{n} x.\frac{M}{x}.\frac{{}^{M-1}C_{x-1}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}}$$

$$= \frac{M}{{}^{N}C_{n}} \sum_{x=1}^{n} \left({}^{M-1}C_{x-1}.{}^{N-M}C_{n-x} \right)$$

$$= \frac{M}{{}^{N}C_{n}} \left[{}^{M-1}C_{0}.{}^{N-M}C_{n-1} + {}^{M-1}C_{1}.{}^{N-M}C_{n-2} + \ldots + {}^{M-1}C_{n-1}.{}^{N-M}C_{0} \right]$$

$$= \frac{M}{{}^{N}C_{n}} \left({}^{N-1}C_{n-1} \right)$$

[This result is obtained using properties of binomial coefficients and involves lot of calculations and hence its derivation may be skipped. It may be noticed that in this result the left upper suffix and also the right lower suffix is the sum of the corresponding suffices of the binomial coefficients involved in each product term. However, the result used in the above expression is enrectangled below for the interesting learners.]

We know that

$$(1+x)^{m+n} = (1+x)^{m}.(1+x)^{n} \qquad \text{[By the method of indices]}$$

Expanding using binomial theorem as explained in Unit 9 of this course, we have

$${}^{m+n}C_{0}.\,x^{m+n} + {}^{m+n}C_{1}.x^{m+n-1} + {}^{m+n}C_{2}.x^{m+n-2} + \ldots + {}^{m+n}C_{m+n}$$

$$= \left({}^{m}C_{0}x^{m} + {}^{m}C_{1}\,x^{m-1} + {}^{m}C_{2}\,x^{m-2} + \ldots + {}^{m}C_{m} \right)$$

$$.\left({}^{n}C_{0}x^{n} + {}^{n}C_{1}\,x^{n-1} + {}^{n}C_{2}\,x^{n-2} + \ldots + {}^{n}C_{n} \right)$$

Comparing coefficients of X^{m+n-r}, we have

$${}^{m+n}C_{r} = \left({}^{m}C_{0}.{}^{n}C_{r} + {}^{m}C_{1}.{}^{n}C_{r-1} + \ldots + {}^{m}C_{r}.{}^{n}C_{0} \right)$$

$$= \frac{M\lfloor n \lfloor N-n}{\lfloor N}.\frac{\lfloor N-1}{\lfloor N-n \lfloor n-1}$$

$$= \frac{M.n\lfloor n-1}{N.\lfloor N-1}.\frac{\lfloor N-1}{\lfloor n-1} = \frac{nM}{N}.$$

$$E\left(X^{2}\right) = E\left[X\left(X-1\right)+X\right]$$

$$= E\left[X\left(X-1\right)\right] + E\left(X\right)$$

$$= \left[\sum_{x=0}^{n} x\left(x-1\right).\frac{{}^{M}C_{x}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}} \right] + \left(\frac{nM}{N} \right)$$

$$= \sum_{x=0}^{n} \left[x\left(x-1\right).\frac{M}{x}.\frac{M-1}{x-1}.\frac{{}^{M-2}C_{x-2}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}} \right] + \left(\frac{nM}{N} \right)$$

$$= \frac{M(M-1)}{{}^{N}C_n}\left[\sum_{x=0}^{n}\left({}^{M-2}C_{x-2}.{}^{N-M}C_{n-x}\right)\right]+\left(\frac{nM}{N}\right)$$

$$= \frac{M(M-1)}{{}^{N}C_n}\left({}^{N-2}C_{n-2}\right)+\left(\frac{nM}{N}\right)$$

[The result in the first term has been obtained using a property of binomial coefficients as done above for finding E(X).]

$$= \frac{M(M-1)\lfloor N-n \lfloor n}{\lfloor N}.\frac{\lfloor N-2}{\lfloor n-2 \lfloor N-n}+\frac{nM}{N}$$

$$= \frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N}$$

Thus,

$$V(X)=E(X^2)-[E(X)]^2=\frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N}-\left(\frac{nM}{N}\right)^2$$

$$= \frac{NM(N-M)(N-n)}{N^2(N-1)} \qquad \text{[On simplification]}$$

Example 2: A jury of 5 members is drawn at random from a voters' list of 100 persons, out of which 60 are non-graduates and 40 are graduates. What is the probability that the jury will consist of 3 graduates?

Solution: The computation of the actual probability is hypergeometric, which is shown as follows:

P **[**2 non-graduates and 3 graduates**]** $= \frac{{}^{60}C_2.{}^{40}C_3}{{}^{100}C_5}$

$$= \frac{60\times59\times40\times39\times38\times5\times4\times3\times2}{2\times6\times100\times99\times98\times97\times96}$$

$$= 0.2323$$

Example 3: Let us suppose that in a lake there are N fish. A catch of 500 fish (all at the same time) is made and these fish are returned alive into the lake after making each with a red spot. After two days, assuming that during this time these 'marked' fish have been distributed themselves 'at random' in the lake and there is no change in the total number of fish, a fresh catch of 400 fish (again, all at once) is made. What is the probability that of these 400 fish, 100 will be having red spots.

Solution: The computation of the probability is hypergeometric and is shown as follows: As marked fish in the lake are 500 and other are $N-500$,

$\therefore$ P[100 marked fish and 300 others] = $\frac{{}^{500}C_{100}.{}^{N-500}C_{300}}{{}^{N}C_{400}}$.

We cannot numerically evaluate this if N is not given. Though N can be estimated using method of Maximum likelihood estimation which you will read in Unit 2 of MST-004 We are not going to estimate it. You may try it as an exercise after reading Unit 2 of MST-004.

Here, let us take an assumed value of N say 5000.

Then,

$$P[X = 100] = \frac{{}^{500}C_{100} \cdot {}^{4500}C_{300}}{{}^{5000}C_{400}}$$

You will agree that the exact computation of this probability is complicated. Such problem is normally there with the use of hypergeometric distribution, especially, if N and M are large. However, if n is small compared to N i.e. if n is such that $\frac{n}{N} < 0.05$, say then there is not much difference between sampling with and without replacement and hence in such cases, the probability obtained by binomial distribution comes out to be approximately equal to that obtained using hypergeometric distribution.

You may now try the following exercise.

E2) A lot of 25 units contains 10 defective units. An engineer inspects 2 randomly selected units from the lot. He/She accepts the lot if both the units are found in good condition, otherwise all the remaining units are inspected. Find the probability that the lot is accepted without further inspection.

We now conclude this unit by giving a summary of what we have covered in it.

11.4 SUMMARY

The following main points have been covered in this unit:

1) A random variable X is said to have a **discrete uniform (rectangular)** distribution if it takes any positive integer value from 1 to n, and its probability mass function is given by

$$P[X = x] = \begin{cases} \frac{1}{n} & \text{for } x = 1, 2, ..., n \\ 0, & \text{otherwise.} \end{cases}$$

where n is called the parameter of the distribution.

2) For **discrete uniform** distribution, **mean** = $\frac{n+1}{2}$ and **variance** = $\frac{n^2 - 1}{12}$.

3) A random variable X is said to follow the **hypergeometic distribution** with parameters N, M and n if it assumes only non-negative integer values and its probability mass function is given by

$$P(X = x) = \begin{cases} \frac{{}^{M}C_{x} \cdot {}^{N-M}C_{n-x}}{{}^{N}C_{n}} & \text{for } x = 0, 1, 2, ..., \min\{n, M\} \\ 0, & \text{otherwise} \end{cases}$$

where n, M, N are positive integers such that $n \le N$, $M \le N$.

4) For **hypergeometric** distribution, **mean** $= \frac{nM}{N}$ and

variance $= \frac{NM(N-M)(N-n)}{N^2(N-1)}$.

11.5 SOLUTIONS/ANSWERS

E1) Let X be the number on the ticket drawn randomly from an urn containing tickets numbered from 1 to 10.

$\therefore$ X is a discrete uniform random variable having the values

1, 2, 3, 4, …, 10 with probability of each of these values equal to $\frac{1}{10}$.

Thus, the expected frequencies for the values of X are obtained as in the following table:

X	$P(X=x)$	Expected/Theoretical frequency $f(x) = N.P[X=x]$ $=150.P[X=x]$
1	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
2	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
3	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
4	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
5	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
6	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
7	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
8	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
9	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
10	$\frac{1}{10}$	$150\times\frac{1}{10}=15$

E2) Here N = 25, M = 10 and n = 2.

$$\text{The desired probability} = P\begin{bmatrix}\text{none of the 2 randomly selected} \\ \text{units is found defective}\end{bmatrix}$$

$$= \frac{{}^{10}C_0 \cdot {}^{25-10}C_2}{{}^{25}C_2} = \frac{(1).{}^{15}C_2}{{}^{25}C_2} = \frac{15\times14}{25\times24} = 0.35.$$

UNIT 12 GEOMETRIC AND NEGATIVE BINOMIAL DISTRIBUTIONS

Structure

12.1 INTRODUCTION

In Units 9 and 11, we have studied the discrete distributions – Bernoulli, Binomial, Discrete Uniform and Hypergeometric. In each of these distributions, the random variable takes finite number of values. There may also be situations where the discrete random variable assumes countably infinite values. Poisson distribution, wherein discrete random variable takes an indefinite number of values with very low probability of occurrence of event, has already been discussed in Unit 10. Dealing with some more situations where discrete random variable assumes countably infinite values, we, in the present unit, discuss geometric and negative binomial distributions. It is pertinent to mention here that negative binomial distribution is a generalization of geometric distribution. Some instances where these distributions can be applied are “deaths of insects”, “number of insect bites”.

Like binomial distribution, geometric and negative binomial distributions also have independent trials with constant probability of success in each trial. But, in binomial distribution, the number of trials (n) is fixed whereas in geometric distribution, trials are performed till first success and in negative binomial distribution trials are performed till a certain number of successes.

Secs. 12.2 and 12.3 of this unit discuss geometric and negative binomial distribution, respectively along with their properties.

Objectives

After studing this unit, you would be able to:

- define the geometric and negative binomial distributions;
- calculate the mean and variance of these distributions;
- compute probabilities of events associated with these distributions;
- identify the situations where these distributions can be applied; and
- know about distinguishing features of these distributions like memoryless property of geometric distribution.

12.2 GEOMETRIC DISTRIBUTION

Let us consider Bernoulli trials i.e. independent trials having the constant probability 'p' of success in each trial. Each trial has two possible outcomes – success or failure. Now, suppose the trial is performed repeatedly till we get the success. Let X be the number of failures preceding the first success. Example of such a situation is "tossing a coin until head turns up". X defined above may take the values 0, 1, 2, …. Letting q be the probability of failure in each trial, we have

$$P\left[X=0\right] = P[\text{Zero failure preceding the first success}]$$
$$= P(S)$$
$$= p,$$

$$P\left[X=1\right] = P[\text{One failure preceding the first success}]$$
$$= P[F \cap S]$$
$$= P(F)\,P(S) \;\; [\because \text{trials are independent}]$$
$$= qp$$

$$P[X = 2] = P[\text{Two failures preceding the first success}]$$
$$= P[F \cap F \cap S]$$
$$= P(F)\,P(F)\,P(S)$$
$$= qqp$$
$$= q^2\,p$$

and so on.

Therefore, in general, probability of x failures preceding the first success is

$$P[X = x] = q^x p;\;\; x = 0, 1, 2, 3, ...$$

Notice that for $x = 0, 1, 2, 3, ...$ the respective probabilities p, qp, q^2p, q^3p,… are the terms of geometric progression series with common ratio q. That is why, the above probability distribution is known as geometric distribution [see Unit 3 of MST-001].

Hence, the above discussion leads to the following definition:

Definition: A random variable X is said to follow geometric distribution if it assumes non-negative integer values and its probability mass function is given by

$$P\left[X=x\right] = \begin{cases} q^x p \;\text{ for } x = 0, 1, 2, ... \\ 0, \;\; \text{otherwise} \end{cases}$$

Notice that

$$\sum_{x=0}^{\infty} q^x p = p + q\,p + q^2 p + q^3 p + ...$$
$$= p[1 + q + q^2 + q^3 + \dots]$$

$$= p\left(\frac{1}{1-q}\right) = \frac{p}{p} = 1$$

[$\because$ sum of infinite terms of G.P. $= \frac{a}{1-r} = \frac{1}{1-q}$ (see Unit 3 of MST-001)]

Now, let us take up some examples of this distribution.

Example 1: An unbiased die is cast until 6 appear. What is the probability that it must be cast more than five times?

Solution: Let p be the probability of a success i.e. getting 6 in a throw of the die

$\therefore$ $p = \frac{1}{6}$ and $q = 1 - p = \frac{5}{6}$

Let X be the number of failures preceding the first success.

$\therefore$ by geometric distribution,

$$P[X = x] = q^x p;\ x = 0, 1, 2, 3, \ldots$$

$$= \left(\frac{5}{6}\right)^x \left(\frac{1}{6}\right) \text{ for } x = 0, 1, 2, 3, \ldots$$

Thus, the desired probability = P[The die is to be cast more than five times]

= P [The number of throws is at least 6]

$$= P\left[\begin{array}{l}\text{The number of failures preceding} \\ \text{the first success is at least 5}\end{array}\right]$$

$$= P[X \geq 5]$$

$$= P[X = 5] + P[X = 6] + P[X = 7] + \ldots$$

$$= \left(\frac{5}{6}\right)^5 \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^6 \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^7 \left(\frac{1}{6}\right) + \ldots$$

$$= \left(\frac{5}{6}\right)^5 \left(\frac{1}{6}\right)\left[1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \ldots\right]$$

$$= \left(\frac{5}{6}\right)^5 \left(\frac{1}{6}\right)\left[\frac{1}{1-\frac{5}{6}}\right] = \left(\frac{5}{6}\right)^5$$

Let us now discuss some properties of geometric distribution.

Mean and Variance

Mean of the geometric distribution is given as

$$\text{Mean} = E(X) = \sum_{x=0}^{\infty} x\, q^x p = p\sum_{x=1}^{\infty} x\, q^x = p\sum_{x=1}^{\infty} x\, q^{x-1}.q$$

$$= pq\sum_{x=1}^{\infty} x\, q^{x-1} = p\,q\sum_{x=1}^{\infty}\frac{d}{dq}\left(q^x\right)$$

$\frac{d}{dq}\left(q^x\right)$ is the differentiation of q^x w.r.t. q where x is kept as constant
[$\because \frac{d}{dx}\left(x^m\right) = mx^{m-1}$, where m is constant (see Unit 6 of MST-001)]

$$= pq\frac{d}{dq}\left[\sum_{x=1}^{\infty} q^x\right] \qquad \left[\begin{array}{l}\because \text{sum of the derivatives is}\\ \text{the derivatives of the sums}\end{array}\right]$$

$$= pq\frac{d}{dq}\left[\sum_{x=1}^{\infty} q^x\right]$$

$$= pq\frac{d}{dq}\left[q + q^2 + q^3 + \ldots\right]$$

$$= pq\frac{d}{dq}\left[\frac{q}{1-q}\right]$$

$$= pq\left[\frac{(1-q) - q(-1)}{(1-q)^2}\right] \qquad \left[\begin{array}{l}\text{Applying quotient rule}\\ \text{of differentiation}\end{array}\right]$$

$$= pq\left[\frac{1-q+q}{p^2}\right]$$

$$= \frac{q}{p}. \qquad \ldots (1)$$

Variance of the geometric distribution is

$V(X) = E(X^2) - [E(X)]^2$,

where

$$E(X^2) = \sum_{x=\infty}^{\infty} x^2 p(x)$$

$$= \sum_{x=0}^{\infty}\left[x(x-1) + x\right]p(x)$$

[$\because x^2 = x(x-1) + x$ (it has already been discussed in Unit 9)]

$$= \sum_{x=0}^{\infty} x(x-1)p(x) + \sum_{x=0}^{\infty} x\,p(x)$$

$$= \sum_{x=2}^{\infty} x(x-1)q^x p + \left(\frac{q}{p}\right) \qquad \text{[Using (1) in second term]}$$

$$= pq^2\sum_{x=2}^{\infty} x(x-1)q^{x-2} + \left(\frac{q}{p}\right) \qquad [\because q^x = q^{x-2}.q^2]$$

$$= pq^2 \sum_{x=2}^{\infty} \frac{d^2}{dq^2}\left(q^x\right) + \frac{q}{p} \quad \left[\begin{array}{l} \because \dfrac{d^2}{dq^2}\left(q^x\right) = \dfrac{d}{dq}\left(\dfrac{d}{dq} q^x\right) \\ = \dfrac{d}{dq}\left(xq^{x-1}\right) = x\left(x-1\right)q^{x-2} \\ \text{treating x as constant} \end{array}\right]$$

$$= pq^2 \frac{d^2}{dq^2}\left(\sum_{x=2}^{\infty} q^x\right) + \frac{q}{p}$$

$$= pq^2 \frac{d^2}{dq^2}\left[q^2 + q^3 + q^4 + \ldots\right] + \frac{q}{p}$$

$$= pq^2 \frac{d^2}{dq^2}\left[\frac{q^2}{1-q}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{\left(1-q\right)2q - q^2\left(-1\right)}{\left(1-q\right)^2}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{2q - 2q^2 + q^2}{\left(1-q\right)^2}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{2q - q^2}{\left(1-q\right)^2}\right] + \frac{q}{p}$$

$$= pq^2 \left[\frac{\left(1-q\right)^2\left(2-2q\right) - \left(2q-q^2\right).2\left(1-q\right)^1\left(-1\right)}{\left(1-q\right)^4}\right] + \frac{q}{p}$$

$$= pq^2 \left[\frac{\left(1-q\right)\left\{2\left(1-q\right)^2 + 2\left(2q-q^2\right)\right\}}{\left(1-q\right)^4}\right] + \frac{q}{p}$$

$$= pq^2 . \frac{p\left[2p^2 + 2q\left(2-q\right)\right]}{p^4} + \frac{q}{p} \qquad \text{as } p = 1 - q$$

$$= 2\left(\frac{q}{p}\right)^2 \left[p^2 + q\left(2-q\right)\right] + \frac{q}{p}$$

$$= 2\left(\frac{q}{p}\right)^2 \left[\left(1-q\right)^2 + q\left(2-q\right)\right] + \frac{q}{p}$$

$$= 2\left(\frac{q}{p}\right)^2 \left[1 + q^2 - 2q + 2q - q^2\right] + \frac{q}{p}$$

$$= 2\frac{q^2}{p^2}\left(1\right) + \frac{q}{p} = \frac{2q^2}{p^2} + \frac{q}{p}$$

$$\therefore \quad V(X) = E(X^2) - [E(X)]^2$$

$$= \frac{2q^2}{p^2} + \frac{q}{p} - \left(\frac{q}{p}\right)^2$$

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

$$= \frac{q}{p} \cdot \frac{1}{p} = \frac{q}{p^2}$$

Remark 1: Variance = $\frac{q}{p^2} = \frac{q}{p.p} = \frac{\text{Mean}}{p}$

$\Rightarrow$ Variance > Mean $\qquad$ $\left[\because p < 1 \Rightarrow \frac{\text{Mean}}{p} > \text{Mean}\right]$

Hence, unlike binomial distribution, variance of the geometric distribution is greater than mean.

Example 2: Comment on the following:

The mean and variance of geometric distribution are 4 and 3 respectively.

Solution: If the given geometric distribution has parameter p(probability of success in each trial).

Then,

Mean = $\frac{q}{p} = 4$ and Variance = $\frac{q}{p^2} = 3$

$\Rightarrow \frac{1}{p} = \frac{3}{4}$

$\Rightarrow p = \frac{4}{3}$, which is impossible, since probability can never exceed unity.

Hence, the given statement is wrong.

Now, you can try the following exercises.

E1) Probability of hitting a target in any attempt is 0.6, what is the probability that it would be hit on fifth attempt?

E2) Determine the geometric distribution for which the mean is 3 and variance is 4.

Lack of Memory Property

Now, let us discuss the distinguishing property of the geometric distribution i.e. the 'lack of memory' property or 'forgetfulness property'. For example, in a random experiment satisfying geometric distribution the wait up to 3 trials (say) for the first success does not affect the probability that one will have to wait for a further 5 trials if it is given that the first two trials are failures. The geometric distribution is the only discrete distribution which has the forgetfulness (memoryless) property. However, there is one continuous distribution which also has the memoryless property and that is the exponential distribution which we will study in Unit 15 of MST-003. The exponential distribution is also the only continuous distribution having this property. It is pertinent to mention here that in several aspects, the geometric distribution is discrete analogs of the exponential distribution.

Let us now give mathematical/statistical discussion on 'memoryless property' of geometric distribution.

Suppose an event occurs at one of the trials 1, 2, 3, 4,… and the occurrence time X has a geometric distribution with probability p. Let X be the number of trials preceding to which one has to wait for successful attempt.

Thus, $P\left[X \geq j\right] = P\left[X = j\right] + P\left[X = j+1\right] + ...$

$$= q^j p + q^{j+1} p + q^{j+2} p + ...$$

$$= q^j p\left[1 + q + q^2 + ...\right]$$

$$= q^j p\left[\frac{1}{1-q}\right] = q^j p\left(\frac{1}{p}\right) = q^j$$

Now, let us consider the event $\left[X > j+k\right]$

Now, $P\left[X \geq j+k \mid X \geq j\right]$ means the conditional probability of waiting for at least j + k unsuccessful trials given that we waited for at least j unsuccessful attempts; and is given by

$$P\left[X \geq j+k \mid X \geq j\right] = \frac{P\left[X \geq j+k \mid X \geq j\right]}{P\left[X \geq j\right]}$$

$$= \frac{P\left[(X \geq j+k) \cap \left(X \geq j\right)\right]}{P\left[X \geq j\right]}$$

$$= \frac{P\left[X \geq j+k\right]}{P\left[X \geq j\right]} \qquad [\because X \geq j+k \text{ implies that} \geq j]$$

$$= \frac{q^{j+k}}{q^j} = q^k$$

$$= P\left[X \geq k\right] \qquad \left[\begin{array}{l}\because P\left[X \geq j\right] = q^j \text{ already} \\ \text{obtained in this section}\end{array}\right]$$

So, $P\left[X \geq j+k \mid X \geq j\right] = P\left[X \geq k\right]$

The above result reveals that the conditional probability of at least first j+k trials are unsuccessful before the first success given that at least first j trial were unsuccessful, is the same as the probability that the first k trials were unsuccessful. So, the probability to get first success remains same if we start counting of k unsuccessful trials from anywhere provided all the trials preceding to it are unsuccessful i.e. the future does not depend on past, it depends only on the present. So, the geometric distribution forgets the preceding trials and hence this property is given the name "forgetfulness property" or "Memoryless property" or "lack of memory" property.

12.3 NEGATIVE BINOMIAL DISTRIBUTION

Negative binomial distribution is a generalisation of geometric distribution. Like geometric distribution, variance of this distribution is also greater than its mean. There are many instances including 'deaths of insects' and 'number of insect bites' where negative binomial distribution is employed.

Negative binomial distribution is a generalisation of geometric distribution in the sense that geometric distribution is the distribution of 'number of failures preceding the first success' whereas the negative binomial distribution is the distribution of 'number of failures preceding the r^{th} success'.

Let X be the random variable which denote the number of failures preceding the r^{th} success. Let p be the probability of a success and let x failures are there preceding the r^{th} success and hence for this the number of trials is $x + r$.

Now, $(x+r)^{th}$ trial is success, but the remaining (r – 1) successes in the $x+r-1$ trials can happen in any r – 1 trials out of the $(x+r-1)$ trials. Thus, happening of first (r – 1) successes in $(x+r-1)$ trials follow binomial distribution with 'p' as the probability of success in each trial and thus is given by

$$^{x+r-1}C_{r-1}\, p^{r-1}\, q^{(x+r-1)-(r-1)} = {}^{x+r-1}C_{r-1}\, p^{r-1} q^{x} \text{ , where } q = 1 - p$$

[$\because$ by binomial distribution, the probability of x successes in n trials with p as the probability of success is $^{n}C_{x}\, p^{x}\, q^{n-x}$.]

Therefore,

P[x failures preceding the r^{th} success]

=P[{First (r – 1) successes in $(x+r-1)$ trials} $\cap${success in $(x+r)^{th}$ trial}]

=P[First(r – 1) successes in $(x+r-1)$ trails]. P[success in $(x+r)^{th}$ trial]

$$= \left({}^{x+r-1}C_{r-1}\, p^{r-1}\, q^{x}\right) p$$

$$= {}^{x+r-1}C_{r-1}\, p^{r}\, q^{x}$$

The above discussion leads to the following definition:

Definition: A random variable X is said to follow a negative binomial distribution with parameters r (a positive integer) and p (0 < p < 1) if its probability mass function is given by:

$$P[X = x] = \begin{cases} {}^{x+r-1}C_{r-1}\, p^r q^x & \text{for } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Now, as we know that

${}^nC_r = {}^nC_{n-r}$, [See 'combination' in Unit 4 of MST-001]

$\therefore$ ${}^{x+r-1}C_{r-1}$ can be written as

$${}^{x+r-1}C_{(x+r-1)-(r-1)} = {}^{x+r-1}C_x$$

$$= \frac{\lfloor x+r-1}{\lfloor x \lfloor x+r-1-x} = \frac{\lfloor r+x-1}{\lfloor x \lfloor r-1}$$

$$= \frac{(r+(x-1))(r+(x-2))\ldots(r+1)(r)\lfloor r-1}{\lfloor x \lfloor r-1}$$

$$= \frac{(r+x-1)(r+x-2)\ldots(r+1)r}{\lfloor x}$$

$$= \frac{\{-(-r-x+1)\}\{-(-r-x+2)\}\ldots\{-(-r-1)\}\{-(-r)\}}{\lfloor x}$$

[$\because$ Numerator is product of x terms from r + 0 to $r+(x-1)$ and we have taken common (−1) from each of these x terms in the product.]

$$= (-1)^x \frac{(-r-x+1)(-r-x+2)\ldots(-r-1)(-r)}{\lfloor x}$$

$$= \frac{(-1)^x (-r)(-r-1)(-r-2)\ldots\{-r-(x-1)\}}{\lfloor x}$$

[Writing the terms in the numerator in reverse order]

$$= (-1)^x \binom{-r}{x}$$

> **Note:** The symbol $\binom{n}{x}$ stands for ${}^nC_x$ if n is positive integer and is equal to $\frac{n(n-1)(n-2)\ldots\{n-(x-1)\}}{\lfloor x}$. We may also use the symbol $\binom{n}{x}$ if n is any real but in this case though it does not stand for ${}^nC_x$, yet it is equal to $\frac{n(n-1)(n-2)\ldots\{n-(n-x)\}}{\lfloor x}$.

Hence, the probability distribution of negative binomnial distribution can be expressed in the following form:

$$P[X = x] = (-1)^x \binom{-r}{x} p^r q^x$$

$$= \binom{-r}{x} (-q)^x p^r \text{ for } x = 0, 1, 2, 3, \ldots$$

$$= \binom{-r}{x} (-q)^x (1)^{-r-x} p^r \text{ for } x = 0, 1, 2, \ldots$$

Here, the expression $\binom{-r}{x} (-q)^x (1)^{-r-x}$ is similar to the binomial distribution $\binom{n}{x} p^x q^{n-x}$

$\therefore \binom{-r}{x} (-q)^x (1)^{-r-x}$ is the general term of $[1 + (-q)]^{-r} = (1-q)^{-r}$

> You have already studied in Unit 9 of this Course that $\binom{n}{x} p^x q^{n-x}$ is the general term of $[q+p]^n$.

and hence

$P[X = x] = \binom{-r}{x} (-q)^x (1)^{-r-x} .p^r$ is the general term of $(1-q)^{-r} p^r$

$\therefore$ P[X = 0], P[X = 1], P[X = 2],… are the successive terms of the binomial expansion $(1-q)^{-r} p^r$ and hence the sum of these probabilities

$$= (1-q)^{-r} p^r$$

$$= p^{-r} p^r \quad [\because 1 - q = p]$$

$$= 1,$$

which must be, being a probability distribution.

Also, as the probabilities of the negative binomial distribution for $X = 0, 1, 2, \ldots$ are the successive terms of

$(1-q)^{-r} p^r = (1-q)^{-r} \left(\frac{1}{p}\right)^{-r} = \left[(1-q)\frac{1}{p}\right]^{-r} = \left[\frac{1}{p} + \left(-\frac{q}{p}\right)\right]^{-r}$, which is a binomial expansion with negative index $(-r)$, it is for this reason the probability distribution given above is called the negative binomial distribution.

Mean and Variance

Mean and variance of the negative binomial distribution can be obtained on observing the form of this distribution and comparing it with the binomial distribution as follows:

The probabilities of binomial distribution for X = 0, 1, 2, … are the successive terms of the binomial expansion of $(q + p)^n$ and the mean and variance obtained for the distribution are

Mean = $np = (n)(p)$ i.e. Product of index and second term in (q + p)

Variance = npq = (n) (p) (q) i.e. Product of index, second term in (q + p) and first term in (q + p)

Similarly, the probabilities of negative binomial distribution for X = 0, 1, 2, ...

are the successive term of the expansion of $\left[\frac{1}{p}+\left(-\frac{q}{p}\right)\right]^{-r}$ and thus, its mean and variance are:

Mean = (index) [second term in $\left\{\frac{1}{p}+\left(\frac{-q}{p}\right)\right\}$] = $(-r)\left[\left(-\frac{q}{p}\right)\right]=\frac{rq}{p}$, and

Variance = (index) [second term in $\left\{\frac{1}{p}+\left(-\frac{q}{p}\right)\right\}$] [First term in $\left\{\frac{1}{p}+\left(-\frac{q}{p}\right)\right\}$]

$$= (-r)\left(-\frac{q}{p}\right)\left(\frac{1}{p}\right)$$

$$=\frac{rq}{p^2}.$$

Remark 2

i) If we take r = 1, we have $P[X = x] = pq^x$ for $x = 0, 1, 2, \ldots$ which is geometric probability distribution.

Hence, geometric distribution is a particular case of negative binomial distribution and the latter may be regarded as the generalisation of the former.

ii) Putting r = 1 in the formulas of mean and variance of negative binomial distribution, we have

Mean = $\frac{(1)q}{p}=\frac{q}{p}$, and

Variance = $\frac{(1)(q)}{p^2}=\frac{q}{p^2}$,

which are the mean and variance of geometric distribution.

Example 3: Find the probability that third head turns up in 5 tosses of an unbiased coin.

Solution: It is a negative binomial situation with $p=\frac{1}{2}$, r = 3,

$x + r = 5 \Rightarrow x = 2.$

$\therefore$ by negative binomial distribution, we have

$$P[X = 2] = {}^{x+r-1}C_{r-1}\, p^r q^x$$

$$= {}^{2+3-1}C_{3-1}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = {}^4C_2\left(\frac{1}{2}\right)^5 = \frac{4\times 3}{2}\times\frac{1}{32} = \frac{3}{16}$$

Example 4: Find the probability that a third child in a family is the family's second daughter, assuming the male and female are equally probable.

Solution: It is a negative binomial situation with

$p = \dfrac{1}{2}$ [$\because$ male and female are equally probable]

$r = 2,\ x + r = 3$

$\Rightarrow\ x = 1$

$\therefore$ by negative binomial distribution,

$$P[X = 1] = {}^{x+r-1}C_{r-1}\, p^r\, q^x = {}^{1+2-1}C_{2-1}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = {}^2C_1\left(\frac{1}{2}\right)^3 = 2\times\frac{1}{8} = \frac{1}{4}.$$

Example 5: A proof-reader catches a misprint in a document with probability 0.8. Find the expected number of misprints in the document in which the proof-reader stops after catching the 20^{th} misprint.

Solution: Let X be the number of misprints not caught by the proof-reader and r be the number of misprints caught by him/her. It is a negative binomial situation where we are to obtain the expected (mean) number of misprints in the document i.e. E(X + r). We will first obtain mean number of misprints which could not be caught by the proof-reader i.e. E(X).

Here, p = 0.8 and hence q = 0.2, r = 20.

Now, by negative binomial distribution,

$$E(X) = \frac{rq}{p} = \frac{(20)(0.2)}{(0.8)} = 5$$

Therefore, $E(X + r) = E(X) + r$ = 5 + 20 = 25.

Hence, the expected number of misprints in the document till he catches the 20^{th} misprint is 25.

Now, we are sure that you will be able to solve the following exercises:

E3) Find the probability that fourth five is obtained on the tenth throw of an unbiased die.

E4) An item is produced by a machine in large numbers. The machine is known to produce 10 per cent defectives. A quality control engineer is testing the item randomly. What is the probability that at least 3 items are examined in order to get 2 defectives?

E5) Find the expected number of children in a family which stops producing children after having the second daughter. Assume, the male and female births are equally probable.

We now conclude this unit by giving a summary of what we have covered in it.

12.4 SUMMARY

The following main points have been covered in this unit:

1) A random variable X is said to follow **geometric distribution** if it assumes non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} q^x p & \text{for } x = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}$$

2) For **geometric** distribution, **mean** $= \dfrac{q}{p}$ and **variance** $= \dfrac{q}{p^2}$.

3) A random variable X is said to follow a **negative binomial distribution** with parameters r (a positive integer) and p (0 < p < 1) if its probability mass function is given by:

$$P(X = x) = \begin{cases} {}^{x+r-1}C_{r-1}\, p^r q^x & \text{for } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

4) For **negative binomial** distribution, **mean** $= \dfrac{rq}{p}$ and **variance** $= \dfrac{rq}{p^2}$.

5) For both these distributions, **variance > mean.**

12.5 SOLUTIONS/ANSWERS

E1) Let p be the probability of success i.e. hitting the target in an attempt.

$\therefore\ p = 0.6,\ q = 1 - p = 0.4$.

Let X be the number of unsuccessful attempts preceding the first successful attempt.
$\therefore$ by geometric distribution,

$$P[X = x] = q^x p \quad \text{for } x = 0, 1, 2, \ldots$$

$$= (0.4)^x (0.6) \quad \text{for } x = 0, 1, 2, \ldots$$

Thus, the desired probability = P[hitting the target in fifth attempt]

= P[The number of unsuccessful attempts before the first success is 4]

$= P[X = 4]$

$= (0.4)^4 (0.6) = (0.0256)(0.6) = 0.01536.$

E2) Let p be the probability of success in an attempt, and q = 1 – p

Now, mean = $\dfrac{q}{p} = 3$ and Variance = $\dfrac{q}{p^2} = 4$

$$\Rightarrow \frac{1}{p} = \frac{4}{3} \quad \Rightarrow p = \frac{3}{4} \text{ and hence } q = \frac{1}{4}.$$

Now, let X be the number of failures preceding the first success,

$$\therefore P\left[X = x\right] = q^x p$$

$$= \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right) \text{ for } x = 0, 1, 2, \ldots$$

This is the desired probability distribution.

E3) It is a negative binomial situation with

$$r = 4,\ x + r = 10 \Rightarrow x = 6,\ p = \frac{1}{6} \text{ and hence } q = \frac{5}{6}$$

$$\therefore\ P\left[X = 6\right] = {}^{x+r-1}C_{r-1} p^r q^x$$

$$= {}^{6+4-1}C_{4-1} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^6$$

$$= {}^{9}C_{3}.\frac{625 \times 25}{36 \times 36 \times 36 \times 36 \times 36}$$

$$= \frac{9 \times 8 \times 7 \times 625 \times 25}{6 \times 36 \times 36 \times 36 \times 36 \times 36} = 0.0217$$

E4) It is a negative binomial situation with $r = 2,\ x + r = 3 \Rightarrow x = 1,\ p = 0.1$ and hence q = 0.9.

Now, the required probability $= P\left[X + r \geq 3\right]$

$$= P\left[X \geq 1\right]$$

$$= 1 - P\left[X = 0\right]$$

$$= 1 - \left[{}^{0+r-1}C_{r-1}\, p^r q^0\right]$$

$$= 1 - \left[{}^{0+2-1}C_{2-1}\, (0.1)^2 (0.9)^0\right]$$

$$= 1 - (1)\,(0.01) = 0.99.$$

E5) It is a negative binomial situation with $p = \frac{1}{2},\ q = \frac{1}{2},\ r = 2.$

Let X be the number of boys in the family

$$\therefore E\left(X\right) = \frac{rq}{p} = \frac{(2)\left(\frac{1}{2}\right)}{\frac{1}{2}} = 2.$$

$$\Rightarrow E(X + r) = E(X) + r = 2 + 2 = 4$$

$\therefore$ the required expected value = 4.

UNIT 13 NORMAL DISTRIBUTION

Structure

13.1 INTRODUCTION

In Units 9 to 12, we have studied standard discrete distributions. From this unit onwards, we are going to discuss standard continuous univariate distributions. This unit and the next unit deal with normal distribution. Normal distribution has wide spread applications. It is being used in almost all data-based research in the field of agriculture, trade, business, industry and the society. For instance, normal distribution is a good approximation to the distribution of heights of randomly selected large number of students studying at the same level in a university.

The normal distribution has a unique position in probability theory, and it can be used as approximation to most of the other distributions. Discrete distributions occurring in practice including binomial, Poisson, hypergeometric, etc. already studied in the previous block (Block 3) can also be approximated by normal distribution. You will notice in the subsequent courses that theory of estimation of population parameters and testing of hypotheses on the basis of sample statistics have also been developed using the concept of normal distribution as most of the sampling distributions tend to normality for large samples. Therefore, study of normal distribution is very important.

Due to various properties and applications of the normal distribution, we have covered it in two units – Units 13 and 14. In the present unit, normal distribution is introduced and explained in Sec. 13.2. Chief characteristics of normal distribution are discussed in Sec. 13.3. Secs. 13.4, 13.5 and 13.6 describes the moments, mode, median and mean deviation about mean of the distribution.

Objectives

After studing this unit, you would be able to:

- introduce and explain the normal distribution;

- know the conditions under which binomial and Poisson distributions tend to normal distribution;
- state various characteristics of the normal distribution;
- compute the moments, mode, median and mean deviation about mean of the distribution; and
- solve various practical problems based on the above properties of normal distribution.

13.2 NORMAL DISTRIBUTION

The concept of normal distribution was initially discovered by English mathematician Abraham De Moivre (1667-1754) in 1733. De Moivre obtained this continuous distribution as a limiting case of binomial distribution. His work was further refined by Pierre S. Laplace (1749-1827) in 1774. But the contribution of Laplace remained unnoticed for long till it was given concrete shape by Karl Gauss (1777-1855) who first made reference to it in 1809 as the distribution of errors in Astronomy. That is why the normal distribution is sometimes called Gaussian distribution. Though, normal distribution can be used as approximation to most of the other distributions, here we are going to discuss (without proof) its approximation to (i) binomial distribution and (ii) Poisson distribution.

Normal Distribution as a Limiting Case of Binomial Distribution

Normal distribution is a limiting case of binomial distribution under the following conditions:

i) n, the number of trials, is indefinitely large i.e. $n \to \infty$;

ii) neither p (the probability of success) nor q (the probability of failure) is too close to zero.

Under these conditions, the binomial distribution can be closely associated by a normal distribution with standardized variable given by $Z = \frac{X - np}{\sqrt{npq}}$. The approximation becomes better with increasing n. In practice, the approximation is very good if both np and nq are greater than 5.

For binomial distribution, you have already studied [see Unit 9 of this course] that

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left[npq(q-p)\right]^2}{\left[npq\right]^3} = \frac{(q-p)^2}{npq},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\left[1+3(n-2)pq\right]}{\left[npq\right]^2} = 3 + \frac{1-6pq}{npq},$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}}, \text{ and}$$

$$\gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}.$$

From the above results, it may be noticed that if $n \to \infty$, then moment coefficient of skewness $(\gamma_1) \to 0$ and the moment coefficient of kurtosis i.e. $\beta_2 \to 3$ or $\gamma_2 \to 0$. Hence, as $n \to \infty$, the distribution becomes symmetrical and the curve of the distribution becomes mesokurtic, which is the main feature of normal distribution.

Normal Distribution as a Limiting Case of Poisson Distribution

You have already studied in Unit 10 of this course that Poisson distribution is a limiting case of binomial distribution under the following conditions:

i) n, the number of trials is indefinitely large i.e. $n \to \infty$

ii) p, the constant probability of success for each trial is very small i.e. $p \to 0$.

iii) np is a finite quantity say 'λ'.

As we have discussed above that there is a relation between the binomial and normal distributions. It can, in fact, be shown that the Poisson distribution approaches a normal distribution with standardized variable given by

$Z = \dfrac{X-\lambda}{\sqrt{\lambda}}$ as λ increases indefinitely.

For Poisson distribution, you have already studied in Unit 10 of the course that

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \Rightarrow \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}; \text{ and}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2}{\lambda^2} + \lambda = 3 + \frac{1}{\lambda} \Rightarrow \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}.$$

Like binomial distribution, here in case of Poisson distribution also it may be noticed from the above results that the moment coefficient of skewness $(\gamma_1) \to 0$ and the moment coefficient of kurtosis i.e. $\beta_2 \to 3$ or $\gamma_2 \to 0$ as $\lambda \to \infty$. Hence, as $\lambda \to \infty$, the distribution becomes symmetrical and the curve of the distribution becomes mesokurtic, which is the main feature of normal distribution.

Under the conditions discussed above, a random variable following a binomial distribution or following a Poisson distribution approaches to follow normal distribution, which is defined as follows:

Definition: A continuous random variable X is said to follow normal distribution with parameters μ $(-\infty < \mu < \infty)$ and $\sigma^2(>0)$ if it takes on any real value and its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \ -\infty < x < \infty;$$

which may also be written as

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, -\infty < x < \infty.$$

Remark

i) The probability function represented by $f(x)$ may also be written as $f(x;\, \mu, \sigma^2)$.

ii) If a random variable X follows normal distribution with mean μ and variance σ^2, then we may write, "X is distributed to $N(\mu, \sigma^2)$" and is expressed as $X \sim N(\mu, \sigma^2)$.

iii) No continuous probability function and hence the normal distribution can be used to obtain the probability of occurrence of a particular value of the random variable. This is because such probability is very small, so instead of specifying the probability of taking a particular value by the random variable, we specify the probability of its lying within interval. For detail discussion on the concept, Sec. 5.4 of Unit 5 may be referred to.

iv) If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X - \mu}{\sigma}$ is standard normal variate having mean '0' and variance '1'. The values of mean and variance of standard normal variate are obtained as under, for which properties of expectation and variance are used (see Unit 8 of this course).

Mean of Z i.e. $E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X-\mu)\right]$

$$= \frac{1}{\sigma}\left[E(X) - \mu\right]$$

$$= \frac{1}{\sigma}\left[\mu - \mu\right] = 0 \qquad [\because E(X) = \text{Mean of } X = \mu\,]$$

Variance of Z i.e. $V(Z) = V\left(\frac{X-\mu}{\sigma}\right)$

$$= \frac{1}{\sigma^2}\left[V(X-\mu)\right] = \frac{1}{\sigma^2}\left[V(X)\right]$$

$$= \frac{1}{\sigma^2}\left(\sigma^2\right) \qquad [\because \text{variance of } X \text{ is } \sigma^2]$$

$$= 1.$$

v) The probability density function of standard normal variate $Z = \frac{X-\mu}{\sigma}$ is given by $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty$.

This result can be obtained on replacing $f(x)$ by $\phi(z)$, x by z, μ by 0 and σ by 1 in the probability density function of normal variate X i.e. in

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty$$

vi) The graph of the normal probability function $f(x)$ with respect to x is famous 'bell-shaped' curve. The top of the bell is directly above the mean μ. For large value of σ, the curve tends to flatten out and for small values of σ, it has a sharp peak as shown in (Fig. 13.1):

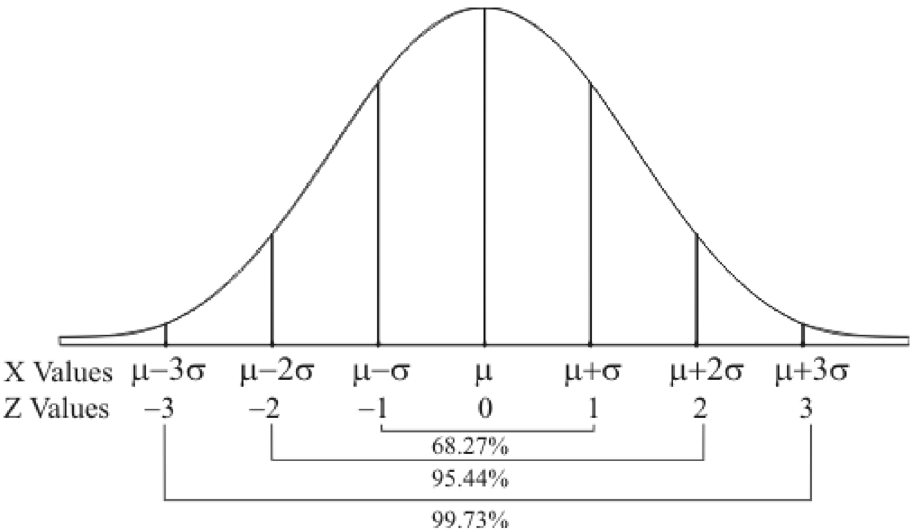


Fig. 13.1

Normal distribution has various properties and large number of applications. It can be used as approximation to most of the other distributions and hence is most important probability distribution in statistical analysis. Theory of estimation of population parameters and testing of hypotheses on the basis of sample statistics (to be discussed in the next course MST-004) have also been developed using the concept of normal distribution as most of the sampling distributions tend to normality for large samples. Normal distribution has become widely and uncritically accepted on the basis of much practical work. As a result, it holds a central position in Statistics.

Let us now take some examples of writing the probability function of normal distribution when mean and variance are specified, and vice-versa:

Example 1: (i) If $X \sim N(40, 25)$ then write down the p.d.f. of X

(ii) If $X \sim N(-36, 20)$ then write down the p.d.f. of X

(iii) If $X \sim N(0, 2)$ then write down the p.d.f. of X

Solution: (i) Here we are given $X \sim N(40, 25)$

$\therefore$ in usual notations, we have

$$\mu = 40,\ \sigma^2 = 25 \Rightarrow \sigma = \pm\sqrt{25}$$

$$\Rightarrow \sigma = 5 \qquad [\because \sigma > 0 \text{ always}]$$

Now, the p.d.f. of random variable X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-40}{5}\right)^2}, \qquad -\infty < x < \infty$$

(ii) Here we are given $X \sim N(-36, 20)$.

$\therefore$ in usual notations, we have

$\mu = -36, \sigma^2 = 20 \Rightarrow \sigma = \sqrt{20}$

Now, the p.d.f. of random variable X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{1}{\sqrt{20}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-(-36)}{\sqrt{20}}\right)^2} = \frac{1}{\sqrt{40\pi}} e^{-\frac{1}{40}(x+36)^2}$$

$$= \frac{1}{2\sqrt{10\pi}} e^{-\frac{1}{40}(x+36)^2}, \qquad -\infty < x < \infty$$

(iii) Here we are given $X \sim N(0, 2)$.

$\therefore$ in usual notations, we have

$\mu = 0, \sigma^2 = 2 \Rightarrow \sigma = \sqrt{2}$

Now, the p.d.f. of random variable X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{\sqrt{2}}\right)^2}$$

$$= \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}x^2}, \qquad -\infty < x < \infty$$

Example 2: Below, in each case, there is given the p.d.f. of a normally distributed random variable. Obtain the parameters (mean and variance) of the variable.

(i) $f(x) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-46}{6}\right)^2}, \qquad -\infty < x < \infty$

(ii) $f(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{32}(x-60)^2}, \qquad -\infty < x < \infty$

Solution: (i) $f(x) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-46}{6}\right)^2}, \qquad -\infty < x < \infty$

Comparing it with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

we have

$\mu = 46, \quad \sigma = 6$

$\therefore$ Mean $= \mu = 46$, variance $= \sigma^2 = 36$

(ii) $f(x) = \dfrac{1}{4\sqrt{2\pi}}\, e^{-\frac{1}{32}(x-60)^2}, \qquad -\infty < x < \infty$

$$= \frac{1}{4\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{X-60}{4}\right)^2}$$

Comparing it with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

we get

$\mu = 60, \quad \sigma = 4$

$\therefore$ Mean $= \mu = 60$, variance $= \sigma^2 = 16$

Here are some exercises for you.

E 1) Write down the p.d.f. of r. v. X in each of the following cases:

(i) $X \sim N\left(\frac{1}{2}, \frac{4}{9}\right)$

(ii) $X \sim N\ (-40, 16)$

E 2) Below, in each case, is given the p.d.f. of a normally distributed random variable. Obtain the parameters (mean and variance) of the variable.

(i) $f(x) = \dfrac{1}{2\sqrt{2\pi}}\, e^{-\frac{x^2}{8}}, \qquad -\infty < x < \infty$

(ii) $f(x) = \dfrac{1}{2\sqrt{\pi}}\, e^{-\frac{1}{4}(x-2)^2}, \qquad -\infty < x < \infty$

Now, we are going to state some important properties of Normal distribution in the next section.

13.3 CHIEF CHARACTERISTICS OF NORMAL DISTRIBUTION

The normal probability distribution with mean μ and variance σ^2 has the following properties:

i) The curve of the normal distribution is bell-shaped as shown in Fig. 13.1 given in Remark (vi) of Sec. 13.2.

ii) The curve of the distribution is completely symmetrical about $x = \mu$ i.e. if we fold the curve at $x = \mu$, both the parts of the curve are the mirror images of each other.

iii) For normal distribution, Mean = Median = Mode

iv) $f(x)$, being the probability, can never be negative and hence no portion of the curve lies below x-axis.

v) Though x-axis becomes closer and closer to the normal curve as the magnitude of the value of x goes towards ∞ or $-\infty$, yet it never touches it.

vi) Normal curve has only one mode.

vii) Central moments of Normal distribution are

$\mu_1 = 0,\ \mu_2 = \sigma^2,\ \mu_3 = 0,\ \mu_4 = 3\sigma^4$ and

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0,\ \beta_2 = \frac{\mu_4}{\mu_2^2} = 3$$

i.e. the distribution is symmetrical and curve is always mesokurtic.

Note: Not only μ_1 and μ_3 are 0 but all the odd order central moments are zero for a normal distribution.

viii) For normal curve,

Q_3 – Median = Median – Q_1

i.e. First and third quartiles of normal distribution are equidistant from median.

ix) Quartile Deviation (Q.D.) = $\frac{Q_3 - Q_1}{2}$ is approximately equal to $\frac{2}{3}$ of the standard deviation.

x) Mean deviation is approximately equal to $\frac{4}{5}$ of the standard deviation.

xi) Q.D. : M.D. : S.D. = $\frac{2}{3}\sigma : \frac{4}{5}\sigma : \sigma = 10 : 12 : 15$

xii) The points of inflexion of the curve are

$$x = \mu \pm \sigma,\ f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}}$$

xiii) If $X_1, X_2, ..., X_n$ are independent normal variables with means $\mu_1, \mu_2, ..., \mu_n$ and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$ respectively then the linear combination $a_1X_1 + a_2X_2 + ... + a_nX_n$ of $X_1, X_2, ..., X_n$ is also a normal variable with
mean $a_1\mu_1 + a_2\mu_2 + ... + a_n\mu_n$ and variance $a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + ... + a_n^2\sigma_n^2$.

xiv) Particularly, sum or difference of two independent normal variates is also a normal variate. If X and Y are two independent normal variates with means μ_1, μ_2 and variances σ_1^2, σ_2^2, then

$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ and $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Also, if $X_1, X_2, \ldots, X_n$ are independent variates each distributed as $N(\mu, \sigma^2)$, then their mean $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

xv) Area property:

$$P[\mu-\sigma < X < \mu+\sigma] = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\,dx = 0.6827,$$

$$\text{Or } P[-1 < Z < 1] = \int_{-1}^{1} \phi(z)\,dz = 0.6827,$$

$$P[\mu-2\sigma < X < \mu+2\sigma] = \int_{\mu-2\sigma}^{\mu+2\sigma} f(x)\,dx = 0.9544,$$

$$\text{Or } P[-2 < Z < 2] = \int_{-2}^{2} \phi(z)\,dz = 0.9544, \text{ and}$$

$$P[\mu-3\sigma < X < \mu+3\sigma] = \int_{\mu-3\sigma}^{\mu+3\sigma} f(x)\,dx = 0.9973.$$

$$\text{Or } P[-3 < Z < 3] = \int_{-3}^{3} \phi(z)\,dz = 0.9973.$$

This property and its applications will be discussed in detail in Unit 14.

Let us now establish some of these properties.

13.4 MOMENTS OF NORMAL DISTRIBUTION

Before finding the moments, following is defined as gamma function [See Unit 16 of the Course also for detail discussion] which is used for computing the even order central moments.

Gamma Function

If n > 0, the integral $\int_0^\infty x^{n-1}e^{-x}dx$ is called a gamma function and is denoted by $\Gamma(n)$.

e.g. $$\int_0^\infty x^2 e^{-x}dx = \int_0^\infty x^{3-1}e^{-x}dx = \Gamma(3)$$

and $$\int_0^\infty x^{-1/2}e^{-x}dx = \int_0^\infty x^{\frac{1}{2}-1}e^{-x}dx = \Gamma\left(\frac{1}{2}\right)$$

Some properties of the gamma function are

i) If n > 1, $\Gamma(n) = (n-1)\Gamma(n-1)$

ii) If n is a positive integer, then $\Gamma n = \underline{|n-1}$

iii) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Now, the first four central moments of normal distribution are obtained as follows:

First Order Central Moment

As first order central moment (μ_1) of any distribution is always zero [see Unit 3 of MST-002], therefore, first order central moment (μ_1) of normal distribution = 0.

Second Order Central Moment

$\mu_2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$ [See Unit 8 of MST-003]

$$= \int_{-\infty}^{\infty} (x-\mu)^2 . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x-\mu = \sigma z$

Differentiating

$\frac{dx}{\sigma} = dz$

$\Rightarrow dx = \sigma dz$

Also, when $x \to -\infty$, we have $z \to -\infty$ and

and when $x \to \infty$, $z \to \infty$

$$\therefore\ \mu_2 = \int_{-\infty}^{\infty} (\sigma z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}z^2} dz$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} z^2 e^{-\frac{z^2}{2}} dz$$

$\because$ on changing z to – z, the integrand i.e. $z^2 e^{-\frac{z^2}{2}}$ does not get changed i.e. it is an even function of z [see Unit 2 of MST-001]. Now, the following property of definite integral can be used:

$\int_{-\infty}^{\infty} f(z)dz = 2\int_{0}^{\infty} f(z)dz$ if $f(z)$ is even function of z

Now, put $z^2 = 2t \Rightarrow z = \sqrt{2}\sqrt{t} = \sqrt{2}\, t^{\frac{1}{2}} \Rightarrow dz = \sqrt{2}\frac{1}{2} t^{-\frac{1}{2}} dt$

$$\therefore\ \mu_2 = \sigma^2\sqrt{\frac{2}{\pi}}\int_0^\infty (2t)e^{-t}\frac{1}{\sqrt{2}t^{\frac{1}{2}}}dt \ = \sigma^2\sqrt{\frac{2}{\pi}}\sqrt{2}\int_0^\infty t^{\frac{1}{2}}e^{-t}dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\int_0^\infty t^{\frac{3}{2}-1}e^{-t}dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) \qquad \text{[By def. of gamma function]}$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\Gamma\frac{1}{2} \qquad \text{[By Property (i) of gamma function]}$$

$$= \frac{\sigma^2}{\sqrt{\pi}}\left(\sqrt{\pi}\right) \qquad \text{[By Property (iii) of gamma function]}$$

$$= \sigma^2$$

Third Order Central Moment

$$\mu_3 = \int_{-\infty}^{\infty}(x-\mu)^3 f(x)dx$$

$$= \int_{-\infty}^{\infty}(x-\mu)^3\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x-\mu = \sigma z \ \Rightarrow\ dx = \sigma dz$

and hence

$$\mu_3 = \int_{-\infty}^{\infty}(\sigma z)^3.\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\sigma dz$$

$$= \sigma^3\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^3 e^{-\frac{1}{2}z^2}dz$$

Now, as integrand $z^3e^{-\frac{1}{2}z^2}$ changes to $-z^3e^{-\frac{1}{2}z^2}$ on changing z to – z i.e. $z^3e^{-\frac{1}{2}z^2}$ is an odd function of z.

Therefore, using the following property of definite integral:

$$\boxed{\int_{-a}^{a} f(z)dz = 0 \ \text{ if f(z) is an odd function of z}}$$

we have,

$$\mu_3 = \sigma^3\frac{1}{\sqrt{2\pi}}(0) \ = 0$$

Fourth Order Central Moment

$$\mu_4 = \int_{-\infty}^{\infty} (x-\mu)^4 f(x)\,dx = \int_{-\infty}^{\infty} (x-\mu)^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Putting $\frac{x-\mu}{\sigma} = z$

$\Rightarrow$ $dx = \sigma dz$

$$\therefore \mu_4 = \int_{-\infty}^{\infty} (\sigma z)^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma^4}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 e^{-\frac{1}{2}z^2} dz = \frac{2\sigma^4}{\sqrt{2\pi}} \int_{0}^{\infty} z^4 e^{-\frac{1}{2}z^2} dz$$

$\because$ integrand $z^4.e^{-\frac{z^2}{2}}$ does not get changed on changing z to – z and hence it is an even function of z [using the same property as used in case of μ_2].

Put $\frac{z^2}{2} = t \Rightarrow z^2 = 2t$

$\Rightarrow 2zdz = 2dt$

$\Rightarrow$ $z\,dz = dt$

$\Rightarrow dz = \frac{dt}{z} = \frac{dt}{\sqrt{2t}}$

$$\therefore \mu_4 = \frac{2\sigma^4}{\sqrt{2\pi}} \int_{0}^{\infty} (2t)^2 e^{-t} \frac{1}{\sqrt{2t}} dt$$

$$= \frac{2\sigma^4.4}{\sqrt{2\pi}\sqrt{2}} \int_{0}^{\infty} t^2 e^{-t} \frac{1}{\sqrt{t}} dt = \frac{4\sigma^4}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{3}{2}} e^{-t} dt$$

$$= \frac{4\sigma^4}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{5}{2}-1} e^{-t} dt$$

$= \frac{4\sigma^4}{\sqrt{\pi}} \Gamma\left(\frac{5}{2}\right)$ [By definition of gamma function]

$= \frac{4\sigma^4}{\sqrt{\pi}} \frac{3}{2} \Gamma\left(\frac{3}{2}\right)$ [By Property (i) of gamma function]

$= \frac{4\sigma^4}{\sqrt{\pi}} \frac{3}{2} \frac{1}{2} \Gamma\left(\frac{1}{2}\right)$ [By Property (i) of gamma function]

$= \frac{3\sigma^4}{\sqrt{\pi}} \sqrt{\pi}$ [Using $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (Property (iii) of gamma function)]

$= 3\sigma^4$

Thus, the first four central moments of normal distribution are

$\mu_1 = 0,\ \mu_2 = \sigma^2,\ \mu_3 = 0,\ \mu_4 = 3\sigma^4.$

$$\Rightarrow \beta_1 = \frac{{\mu_3}^2}{{\mu_2}^3} = 0,\ \beta_2 = \frac{\mu_4}{{\mu_2}^2} = \frac{3\sigma^4}{\left(\sigma^2\right)^2} = \frac{3\sigma^4}{\sigma^4} = 3$$

Therefore, moment coefficient of skewness $(\gamma_1) = 0$

$\Rightarrow$ the distribution is symmetrical.

The moment coefficient of kurtosis is $\beta_2 = 3$ or $\gamma_2 = 0.$

$\Rightarrow$ The curve of the normal distribution is mesokurtic.

Now, let us obtain the mode and median for normal distribution in the next section.

13.5 MODE AND MEDIAN OF NORMAL DISTRIBUTION

Mode

Let $X \sim N\left(\mu, \sigma^2\right)$, then p.d.f. of X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \text{...(1)}$$

$$, -\infty < x < \infty$$

Taking logarithm on both sides of (1), we get

$$\log f(x) = \log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \log e \qquad \left[\begin{array}{l}\because \log mn = \log m + \log n \\ \text{and} \log m^n = n \log m\end{array}\right]$$

$$= \log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}(x-\mu)^2 \qquad \text{as } \log e = 1$$

Differentiating w.r.t.x

$$\frac{1}{f(x)} f'(x) = 0 - \frac{1}{2\sigma^2} 2(x-\mu) = -\frac{(x-\mu)}{\sigma^2}$$

$$\Rightarrow f'(x) = -\frac{(x-\mu)}{\sigma^2} f(x) \qquad \ldots (2)$$

For maximum or minimum

$f'(x) = 0$

$$\Rightarrow -\frac{(x-\mu)}{\sigma^2} f(x) = 0$$

$\Rightarrow x - \mu = 0$ as $f(x) \neq 0$

$\Rightarrow x = \mu$

Now differentiating (2) w.r.t. x, we have

$$f''(x) = -\frac{(x-\mu)}{\sigma} f'(x) - \frac{(1)}{\sigma^2} f(x)$$

$$f''(x)\Big|_{\text{at } x=\mu} = 0 - \frac{f(\mu)}{\sigma^2} = -\frac{f(\mu)}{\sigma^2} < 0$$

$\therefore x = \mu$ is point where function has a maximum value.

$\Rightarrow$ Mode of X is μ.

Median

Let M denote the median of the normally distributed random variable X.

We know that median divides the distribution into two equal parts

$$\therefore \int_{-\infty}^{M} f(x)dx = \int_{M}^{\infty} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

In the first integral, let us put $\dfrac{x-\mu}{\sigma} = z$

Therefore, $dx = \sigma\, dz$

Also when $x = \mu \Rightarrow z = 0$, and

when $x \to -\infty \Rightarrow z \to -\infty$.

Thus, we have

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2} dz + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \frac{1}{2} + \int_{\mu}^{M} f(x)\,dx = \frac{1}{2} \qquad \left[\begin{array}{l} \because Z \text{ is s.n.v. with p.d.f. } \phi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \\ \text{So } \displaystyle\int_{-\infty}^{\infty} \phi(z)dz = 1 \Rightarrow \int_{-\infty}^{0} \phi(z)dz = \frac{1}{2} \end{array}\right]$$

$$\Rightarrow \int_{\mu}^{M} f(x)dx = 0$$

$$\Rightarrow M = \mu \qquad \text{as } f(x) \neq 0$$

$\therefore$ Median of $X = \mu$

From the above two results, we see that

$\boxed{\textbf{Mean = Median = Mode = } \mu}$

13.6 MEAN DEVIATION ABOUT THE MEAN

Mean deviation about mean for normal distribution is

$$= \int_{-\infty}^{\infty} |x - \text{Mean}|\, f(x)dx \qquad \text{[See Section 8.4 of Unit 8]}$$

$$= \int_{-\infty}^{\infty} |x-\mu| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x - \mu = \sigma z$

$$\Rightarrow \frac{dx}{\sigma} = dz$$

$$\therefore \text{ M.D. about mean} = \int_{-\infty}^{\infty} |\sigma z| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-\frac{1}{2}z^2} dz$$

Now, $|z|e^{-\frac{1}{2}z^2}$ (integrand) is an even function z as it does not get changed on changing z to –z, $\therefore$ by the property,

"$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$, if $f(x)$ is an even function of x", we have

$$\text{M.D. about mean} = \frac{\sigma}{\sqrt{2\pi}} 2\int_{0}^{\infty} |z| e^{-\frac{1}{2}z^2} dz$$

Now, as the range of z is from 0 to ∞ i.e. z takes non-negative values,

$\therefore$ $|z|$ = z and hence

$$\text{M.D. about mean} = \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} z e^{-\frac{1}{2}z^2} dz$$

Put $\frac{z^2}{2} = t \Rightarrow z^2 = 2t \Rightarrow 2zdz = 2dt \Rightarrow zdz = dt$

$$\therefore \text{ M.D. about mean} = \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t} dt = \sqrt{2}\frac{\sigma}{\sqrt{\pi}} \left[\frac{e^{-t}}{-1}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\sigma[-0+1] = \sqrt{\frac{2}{\pi}}\sigma$$

In practice, instead of $\sqrt{\frac{2}{\pi}}\sigma,$ its approximate value is mostly used and that is $\frac{4}{5}\sigma.$

$$\because \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2\times 7}{22}} = \sqrt{\frac{7}{11}} = \sqrt{0.6364} = 0.7977 = 0.08 \text{ or } \frac{4}{5}(\text{approx.})$$

Let us now take up some problems based on properties of Normal Distribution in the next section.

13.7 SOME PROBLEMS BASED ON PROPERTIES OF NORMAL DISTRIBUTION

Example 3: If X_1 and X_2 are two independent variates each distributed as N(0, 1), then write the distribution of (i) $X_1 + X_2$. (ii) $X_1 - X_2$.

Solution: We know that, if X_1 and X_2 are two independent normal variates s.t.

$X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ and $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$

then

$X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$, and

$X_1 - X_2 \sim N\left(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2\right)$ [See Property xiii (Section 13.3)]

Here, $X_1 \sim N(0,1),\ X_2 \sim N(0,1)$

$\therefore$ i) $X_1 + X_2 \sim N(0+0,\ 1+1)$

i.e. $X_1 + X_2 \sim N(0,\ 2)$, and

ii) $X_1 - X_2 \sim N(0-0,\ 1+1)$

i.e. $X_1 - X_2 \sim N(0,\ 2)$

Example 4: If $X \sim N(30, 25)$, find the mean deviation about mean.

Solution: Here $\mu = 30$, $\sigma^2 = 25 \Rightarrow \sigma = 5$.

$\therefore$ Mean deviation about mean = $\sqrt{\frac{2}{\pi}}\sigma = \sqrt{\frac{2}{\pi}}.5 = 5\sqrt{\frac{2}{\pi}}$

Example 5: If $X \sim N(0, 1)$, what are its first four central moments?

Solution: Here $\mu = 0$, $\sigma^2 = 1 \Rightarrow \sigma = 1$.

$\therefore$ first four central moments are:

$\mu_1 = 0,\ \mu_2 = \sigma^2 = 1,\ \mu_3 = 0,\ \mu_4 = 3\sigma^4 = 3.$

Example 6: If X_1, X_2 are independent variates such that $X_1 \sim N(40, 25)$, $X_2 \sim N(60, 36)$, then find mean and variance of (i) $X = 2X_1 + 3X_2$
(ii) $Y = 3X_1 - 2X_2$

Solution: Here $X_1 \sim N(40, 25)$, $X_2 \sim N(60, 36)$

$\therefore$ Mean of $X_1 = E(X_1) = 40$

Variance of $X_1 = Var(X_1) = 25$

Mean of $X_2 = E(X_2) = 60$

Variance of $X_2 = Var(X_2) = 36$

Now,

(i) Mean of X = $E(X) = E(2X_1 + 3X_2) = E(2X_1) + E(3X_2)$

$$= 2E(X_1) + 3E(X_2) = 2 \times 40 + 3 \times 60 = 80 + 180 = 260$$

Var (X) = $Var(2X_1 + 3X_2)$

$$= Var(2X_1) + Var(3X_2) \qquad [\because X_1 \text{ and } X_2 \text{ are independent}]$$

$$= 4Var(X_1) + 9Var(X_2)$$

$$= 4 \times 25 + 9 \times 36 = 100 + 324 = 424$$

(ii) Mean of Y = $E(Y) = E(3X_1 - 2X_2)$

$$= E(3X_1) + E(-2X_2)$$

$$= 3E(X_1) + (-2)E(X_2)$$

$$= 3 \times 40 - 2 \times (60) = 120 - 120 = 0$$

Var (Y) = Var $(3X_1 - 2X_2)$

$$= Var(3X_1) + Var(-2X_2)$$

$$= (3)^2 Var(X_1) + (-2)^2 Var(X_2)$$

$$= 9 \times 25 + 4 \times 36 = 225 + 144 = 369$$

You can now try some exercises based on the properties of normal distribution which you have studied in the present unit.

E3) If X_1 and X_2 are two independent normal variates with means 30, 40 and variances 25, 35 respectively. Find the mean and variance of

i) $X_1 + X_2$

ii) $X_1 - X_2$

E4) If $X \sim N(50, 225)$, find its Quartile deviation.

E5) If X_1 and X_2 are independent variates with each distributed as N (50, 64), what is the distribution of $\frac{X_1 + X_2}{2}$?

E6) For a normal distribution, the first moment about 5 is 30 and the fourth moment about 35 is 768. Find the mean and standard deviation of the distribution.

13.8 SUMMARY

The following main points have been covered in this unit:

1) A continuous random variable X is said to follow normal distribution with parameters μ ($-\infty < \mu < \infty$) and σ^2(>0) if it takes on any real value and its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \ -\infty < x < \infty$$

2) If $X \sim N(\mu, \sigma^2)$, then $Z = \dfrac{X-\mu}{\sigma}$ is standard normal variate.

3) The curve of the normal distribution is bell-shaped and is completely symmetrical about $x = \mu$.

4) For normal distribution, Mean = Median = Mode.

5) Q_3 – Median = Median – Q_1

6) Quartile Deviation (Q.D.) = $\dfrac{Q_3 - Q_1}{2}$ is approximately equal to $\dfrac{2}{3}$ of the standard deviation.

7) Mean deviation is approximately equal to $\dfrac{4}{5}$ of the standard deviation.

8) Central moments of Normal distribution are

$$\mu_1 = 0,\ \mu_2 = \sigma^2,\ \mu_3 = 0,\ \mu_4 = 3\sigma^4$$

9) Moment coefficient of skewness is zero and the curve is always mesokurtic.

10) Sum of independent normal variables is also a normal variable.

13.9 SOLUTIONS/ANSWERS

E 1) (i) Here we are given $X \sim N\left(\frac{1}{2}, \frac{4}{9}\right)$

$\therefore$ in usual notations, we have

$$\mu = \frac{1}{2}, \quad \sigma^2 = \frac{4}{9} \quad \Rightarrow \sigma = \frac{2}{3}$$

Now, p.d.f. of r.v. X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

$$= \frac{1}{\frac{2}{3}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1/2}{2/3}\right)^2}$$

$$= \frac{3}{2\sqrt{2\pi}}\, e^{-\frac{9}{2}\left(\frac{2x-1}{4}\right)^2}, \qquad -\infty < x < \infty$$

(ii) Here we are given $X \sim N(-40, 16)$

$\therefore$ in usual notations, we have

$\mu = -40, \quad \sigma^2 = 16 \quad \Rightarrow \sigma = 4$

Now, p.d.f. of r.v. X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

$$= \frac{1}{4\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-(-40)}{4}\right)^2}$$

$$= \frac{1}{4\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x+40}{4}\right)^2}, \quad -\infty < x < \infty$$

E 2) (i) $f(x) = \frac{1}{2\sqrt{2\pi}}\, e^{-\frac{x^2}{8}}, \qquad -\infty < x < \infty$

$$= \frac{1}{2\sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{x^2}{4}}$$

$$= \frac{1}{2\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-0}{2}\right)^2} \qquad \ldots(1) \qquad , -\infty < x < \infty$$

Comparing (1) with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

we get

$\mu = 0, \qquad \sigma = 2$

$\therefore$ Mean $= \mu = 0$ and variance $= \sigma^2 = (2)^2 = 4$

(ii) $f(x) = \frac{1}{2\sqrt{\pi}}\, e^{-\frac{1}{4}(x-2)^2}, \qquad -\infty < x < \infty$

$$= \frac{1}{\sqrt{2}\times\sqrt{2}\sqrt{\pi}}\, e^{-\frac{1}{2\times 2}(x-2)^2}$$

$$= \frac{1}{\sqrt{2}\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-2}{\sqrt{2}}\right)^2} \qquad \ldots(1) \qquad , -\infty < x < \infty$$

Comparing (1) with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad , -\infty < x < \infty$$

we get

$\mu = 2, \quad \sigma = \sqrt{2}$

$\therefore$ Mean $= \mu = 2$ and variance $= \sigma^2 = (\sqrt{2})^2 = 2$

E3) i) $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^{\,2} + \sigma_2^{\,2})$

$\Rightarrow X_1 + X_2 \sim N(30 + 40, 25 + 35)$

$\Rightarrow X_1 + X_2 \sim N(70, 60)$

ii) $X_1 - X_2 \sim N(30 - 40, 25 + 35)$

$\Rightarrow X_1 - X_2 \sim N(-10, 60)$

E4) As $\sigma^2 = 225$

$\Rightarrow \sigma = 15$

and hence Q.D. $= \frac{2}{3}\sigma = \frac{2}{3} \times 15 = 10$

E5) We know that if $X_1, X_2, .., X_n$ are independent variates each distributed as $N(\mu, \sigma^2)$, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Here X_1 and X_2 are independent variates each distributed as N(50, 64),

$\therefore$ their mean i.e. $\bar{X}$ i.e. $\frac{X_1 + X_2}{2} \sim N\left(50, \frac{64}{2}\right)$

i.e $\bar{X} \sim N(50, 32)$.

E6) We know that $\mu_1' = \bar{x} - A$ [See Unit 3 of MST-002]

where μ_1' is the first moment about A.

$\therefore 30 = \bar{x} - 5 \Rightarrow \bar{x} = 35 \Rightarrow$ Mean = 35

Given that fourth moment about 35 is 768. But mean is 35, and hence the fourth moment about mean = 768.

$\Rightarrow \mu_4 = 768$

$\Rightarrow 3\sigma^4 = 768$

$\Rightarrow \sigma^4 = \frac{768}{3} \qquad [\because \mu_4 = 3\sigma^4]$

$\Rightarrow \sigma^4 = 256 = (4)^4 \Rightarrow \sigma = 4.$

UNIT 14 AREA PROPERTY OF NORMAL DISTRIBUTION

Structure

14.1 INTRODUCTION

In Unit 13, you have studied normal distribution and its chief characteristics. Some characteristics including moments, mode, median, mean deviation about mean have been established too in Unit 13. The area property of normal distribution has just been touched in the preceding unit. Area property is very important property and has lot of applications and hence it needs to be studied in detail. Hence, in the Unit 14 this property with its diversified applications has been discussed in detail. Fitting of normal distribution to the observed data and computation of expected frequencies have also been discussed in one of the sections i.e. Sec. 14.3 of this unit.

Objectives

After studing this unit, you would be able to:

- describe the importance of area property of normal distribution;
- explain use of the area property to solve many practical life problems; and
- fit a normal distribution to the observed data and compute the expected frequencies using area property.

14.2 AREA PROPERTY OF NORMAL DISTRIBUTION

Let X be a normal variate having the mean μ and variance σ^2.

Suppose we are interested in finding $P[\mu < X < x_1]$ [See Fig.14.1]

Now, $P[\mu < X < x_1] = \int_{\mu}^{x_1} f(x)\,dx = \int_{\mu}^{x_1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

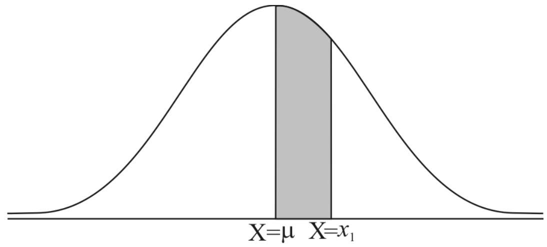


Fig. 14.1: P[μ < X < x_1]

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x - \mu = \sigma z$

$$\frac{dx}{\sigma} = dz$$

Also, when X = μ, Z = 0

and when $X = x_1,\ Z = \frac{x_1 - \mu}{\sigma} = z_1(\text{say})$

$$\therefore\ P[\mu < X < x_1] = P[0 < Z < z_1] = \int_0^{z_1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \sigma dz$$

$$= \int_0^{z_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} dz = \int_0^{z_1} \phi(z)\, dz$$

where $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ is the probability density function of standard normal variate and the definite integral $\int_0^{z_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$ i.e. $\int_0^{z_1} \phi(z) dz$ represents the area under standard normal curve between the ordinates at Z = 0 and Z = z_1. (Fig. 14.2).

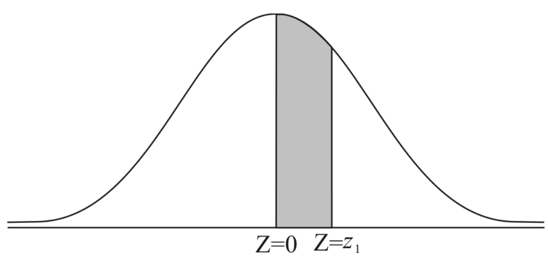


Fig. 14.2: P[0 < Z < z_1]

You need not to evaluate the integral to find the area. Table is available to find such area for different values of z_1.

Here, we have transformed the integral from

$$\int_{\mu}^{x_1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \text{ to } \int_0^{z_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

i.e. we have transformed normal variate 'X' to standard normal variate (S.N.V.) $Z = \frac{X-\mu}{\sigma}$.

This is because, the computation of

$\int_{\mu}^{x_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$ requires construction of separate tables for different values of μ and σ as the normal variate X may have any values of mean and standard deviation and hence different tables are required for different μ and σ. So, infinitely many tables are required to be constructed which is impossible. But beauty of standard normal variate is that its mean is always '0' and standard deviation is always '1' as shown in Unit 13. So, whatever the values of mean and standard deviation of a normal variate be, the mean and standard deviation on transforming it to the standard normal variate are always '0' and '1' respectively and hence only one table is required.

In particular,

$$P[\mu-\sigma < X < \mu+\sigma] = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\,dx \qquad [\text{See Fig.14.3}]$$

$$\Rightarrow P[-1 < Z < 1] = \int_{-1}^{1} \phi(z)\,dz \left[\begin{array}{l} \because Z = \frac{X-\mu}{\sigma} \text{ when } X = \mu-\sigma, Z = \frac{\mu-\sigma-\mu}{\sigma} = -1 \\ \text{when } X = \mu+\sigma, Z = \frac{\mu+\sigma-\mu}{\sigma} = \frac{\sigma}{\sigma} = 1 \end{array}\right]$$

$$= 2\int_{0}^{1} \phi(z)\,dz \qquad [\text{By Symmetry}]$$

$$= 2 \times 0.34135 \qquad \left[\begin{array}{l}\text{From the table given in the} \\ \text{Appendix at the end of the unit}\end{array}\right]$$

$$= 0.6827$$

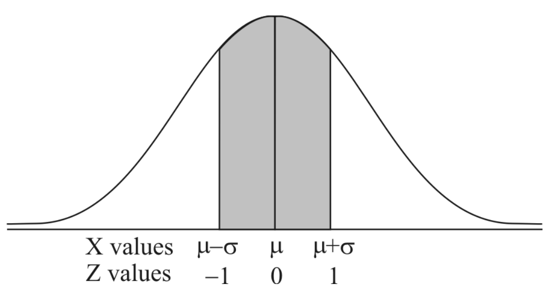


Fig. 14.3: Area within the Range $\mu \pm \sigma$

Similarly,

$$P[\mu-2\sigma < X < \mu+2\sigma] = \int_{\mu-2\sigma}^{\mu+2\sigma} f(x)\,dx \qquad [\text{See Fig.14.4}]$$

$$\Rightarrow \quad P[-2 < Z < 2] = \int_{-2}^{2} \phi(z)\,dx \;\; = 2\int_{0}^{2} \phi(z)\,dx \quad \left[\begin{array}{l} \because \text{for } Z = \dfrac{X-\mu}{\sigma}, \text{ we have} \\ Z = -2 \text{ when} X = \mu - 2\sigma \\ \text{and } Z = 2 \text{ when } X = \mu + 2\sigma \end{array}\right]$$

$$= 2 \times 0.4772 \quad \left[\begin{array}{l} \text{From the table given in the} \\ \text{Appendix at the end of the unit} \end{array}\right]$$

$$= 0.9544$$

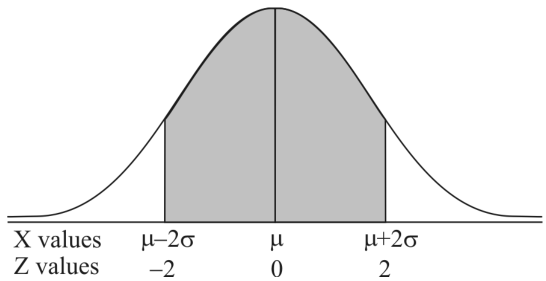


Fig. 14.4: Area within the Range $\mu \pm 2\sigma$

and

$$P[\mu - 3\sigma < X < \mu + 3\sigma] = P[-3 < Z < 3] = 2.P\ [0 < Z < 3] \quad \text{[See Fig. 14.5]}$$

$$= 2 \times 0.49865 = 0.9973$$

$\therefore$ P[X lies within the range $\mu \pm 3\sigma$] = 0.9973

$\Rightarrow$ P[X lies outside the range $\mu \pm 3\sigma$] = 1 – 0.9973 = 0.0027

which is very small and hence usually we expect a normal variate to lie within the range from – 3 to 3, though, theoretically it ranges from $-\infty$ to ∞.

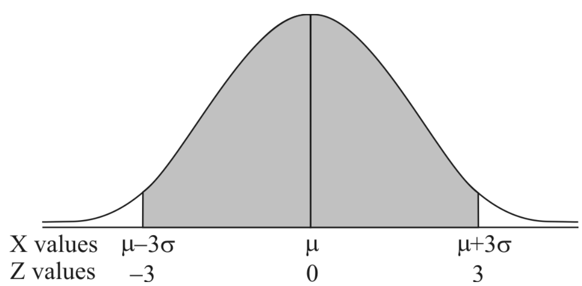


Fig. 14.5: Area within the Range $\mu \pm 3\sigma$

From the above discussion, we conclude that while solving numerical problems, we need to transform the given normal variate into standard normal variate because tables for the area under every normal curve, being infinitely many, cannot be made available whereas the standard normal curve is one and hence table for area under this curve can be made available and this is given in the Appendix at the end of this unit.

Example 1: If X ~ N (45, 16) and Z is the standard normal variable (S.N.V.) i.e $Z = \dfrac{X-\mu}{\sigma}$ then find Z scores corresponding to the following values of X.

(i) X = 45 (ii) X = 53 (iii) X = 41 (iv) X = 47

Solution: We are given X ~ N (45, 16)

∴ In usual notations, we have

$\mu = 45,\ \sigma^2 = 16 \Rightarrow \sigma = \pm\sqrt{16} \Rightarrow \sigma = 4 \quad [\because \sigma > 0 \text{ always}]$

Now $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-45}{4}$

(i) When X = 45, $Z = \dfrac{45-45}{4} = \dfrac{0}{4} = 0$

(ii) When X = 53, $Z = \dfrac{53-45}{4} = \dfrac{8}{4} = 2$

(iii) When X = 41, $Z = \dfrac{41-45}{4} = \dfrac{-4}{4} = -1$

(iv) When X = 47, $Z = \dfrac{47-45}{4} = \dfrac{2}{4} = 0.5$

Example 2: If the r.v. X is normally distributed with mean 80 and standard deviation 5, then find

(i) $P[X > 95]$, (ii) $P[X < 72]$, (iii) $P[60.5 < X < 90]$,

(iv) $P[85 < X < 97]$, and (v) $P[64 < X < 76]$

Solution: Here we are given that X is normally distributed with mean 80 and standard deviation (S.D.) 5.

i.e. $\text{Mean} = \mu = 80$ and $\text{variance} = \sigma^2 = (\text{S.D.})^2 = 25.$

If Z is the S.N.V., then $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-80}{5}$

Now

(i) X = 95, $Z = \dfrac{95-80}{5} = \dfrac{15}{5} = 3$

$\therefore P[X > 95] = P[Z > 3]$ [See Fig.14.6]

$= 0.5 - P[0 < Z < 3]$

= 0.5 – 0.4987 [Using table area under normal curve]

= 0.0013

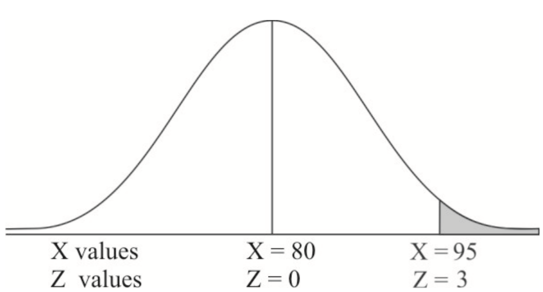


Fig. 14.6: Area to the Right of X= 95

(ii) $X = 72,\ Z = \dfrac{72-80}{5} = \dfrac{-8}{5} = -1.6$

$\therefore P[X < 72] = P[Z < -1.6]$ [See Fig.14.7]

$= P[Z > 1.6]$ $\left[\begin{array}{l}\because \text{normal curve is symmetrical} \\ \text{about the line } Z = 0\end{array}\right]$

$= 0.5 - P[0 < Z < 1.6]$

$= 0.5 - 0.4452$ [Using table area under normal curve]

$= 0.0548$

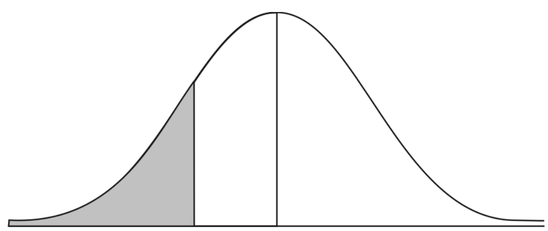


Fig. 14.7: Area to the Left of X = 72

(iii) $X = 60.5,\quad Z = \dfrac{60.5-80}{5} = \dfrac{-19.5}{5} = -3.9$

$X = 90,\quad Z = \dfrac{90-80}{5} = \dfrac{10}{5} = 2$

$\therefore P[60.5 < X < 90] = P[-3.9 < X < 2]$ [See Fig.14.8]

$= P[-3.9 < X < 0] + P[0 < Z < 2]$

$= P[0 < X < 3.9] + P[0 < Z < 2]$ $\left[\begin{array}{l}\because \text{normal curve is} \\ \text{symmetrical about} \\ \text{the line } Z = 0\end{array}\right]$

$= 0.5000 + 0.4772$ $\left[\begin{array}{l}\text{Using table area} \\ \text{under normal curve}\end{array}\right]$

$= 0.9772$

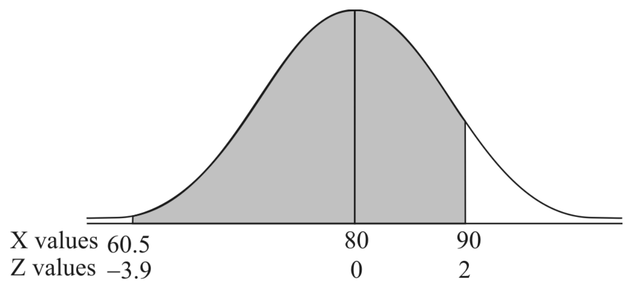


Fig. 14.8: Area between X = 60.5 and X = 90

(iv) $X = 85, \quad Z = \dfrac{85-80}{5} = \dfrac{5}{5} = 1$

$X = 97, Z = \dfrac{97-80}{5} = \dfrac{17}{5} = 3.4$

$\therefore P[85 < X < 97] = P[1 < Z < 3.4] \quad$ [See Fig.14.9]

$= P[0 < Z < 3.4] - P[0 < Z < 1]$

= 0.4997 – 0.3413 [Using table area under normal curve]

= 0.1584

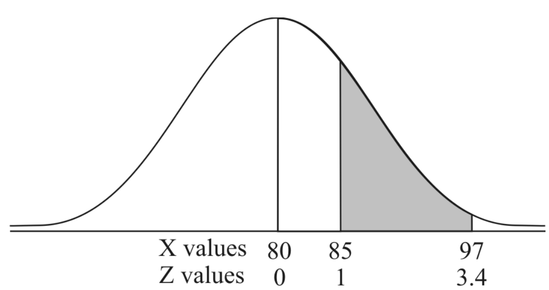


Fig. 1 4.9: Area between X = 85 and X= 97

(v) $X = 64, \quad Z = \dfrac{64-80}{5} = \dfrac{-16}{5} = -3.2$

$X = 76, \quad Z = \dfrac{76-80}{5} = \dfrac{-4}{5} = -0.8$

$P[64 < X < 76] = P[-3.2 < Z < -0.8] \quad$ [See Fig.14.10]

$= P[0.8 < Z < 3.2] \quad \left[\begin{array}{l}\because \text{normal curve is symmetrical} \\ \text{about the line } Z = 0\end{array}\right]$

$= P[0 < Z < 3.2] - P[0 < Z < 0.8]$

$= 0.4993 - 0.2881 \quad \left[\begin{array}{l}\text{Using table area} \\ \text{under normal curve}\end{array}\right]$

= 0.2112

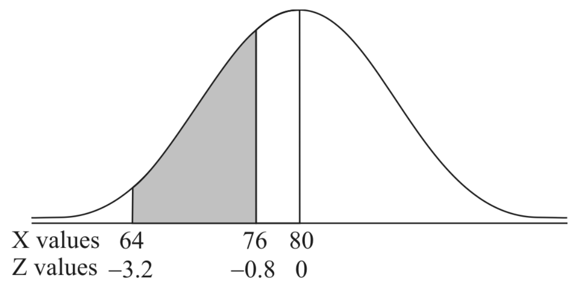


Fig. 14.10: Area between X = 64 and X= 76

Example 3: In a university the mean weight of 1000 male students is 60 kg and standard deviation is 16 kg.

(a) Find the number of male students having their weights
  i) less than 55 kg
  ii) more than 70 kg
  iii) between 45 kg and 65 kg

(b) What is the lowest weight of the 100 heaviest male students?

(Assuming that the weights are normally distributed)

Solution: Let X be a normal variate, "The weights of the male students of the university". Here, we are given that μ = 60 kg, σ = 16 kg, therefore,

$X \sim N(60, 256)$.

We know that if $X \sim N(\mu, \sigma^2)$, then the standard normal variate is given by $Z = \frac{X-\mu}{\sigma}$.

Hence, for the given information, $Z = \frac{X-60}{16}$

(a) i) For X = 55, $Z = \frac{55-60}{16} = -0.3125 \simeq -0.31$.

Therefore,

$P[X < 55] = P[Z < -0.31] = P[Z > 0.31]$ [See Fig. 14.11]

$= 0.5 - P[0 < Z < 0.31]$ $\left[\begin{array}{l}\because \text{area on both} \\ \text{sides of } Z = 0 \text{ is } 0.5\end{array}\right]$

$= 0.5 - 0.1217$ $\left[\begin{array}{l}\text{Using table area} \\ \text{under normal curve}\end{array}\right]$

$= 0.3783$

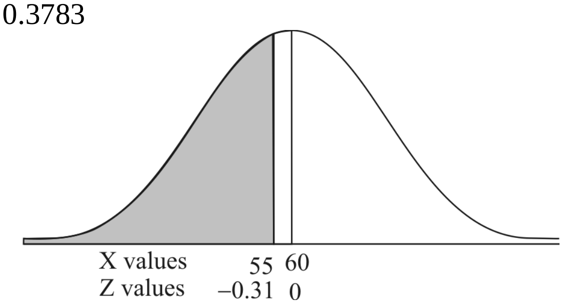


Fig. 14.11: Area Representing Students having Less than 55 kg weight

Number of male students having weight less than 55 kg = $N \times P(X < 55)$

$= 1000 \times 0.3783$

$= 378$

ii) For X = 70, $Z = \dfrac{70-60}{16} = 0.625 \approx 0.63$

$P[X>70] = P\,[Z > 0.63]$ [See Fig. 14.12]

$= 0.5 - P\,[0 < Z < 0.63]$ $\left[\begin{array}{l}\because \text{area on both} \\ \text{sides of } Z = 0 \text{ is } 0.5\end{array}\right]$

$= 0.5 - 0.2357$ $\left[\begin{array}{l}\text{Using table area} \\ \text{under normal curve}\end{array}\right]$

$= 0.2643$

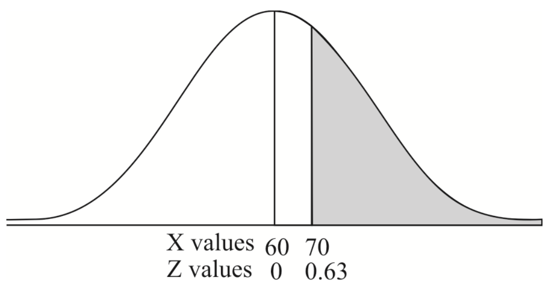


Fig. 14.12: Area Representing Students having More than 70 kg weight

Number of male students having weight more than 70 kg = $N \times P[X > 70]$

$= 1000 \times 0.2643$

$= 264$

iii) For X = 45, $Z = \dfrac{45-60}{16} = -0.9375 \approx -0.94$

For X = 65, $Z = \dfrac{65-60}{16} = 0.3125 \approx 0.31$

$P\left[45 < X < 65\right] = P\left[-0.94 < Z < 0.31\right]$ [See Fig. 14.13]

$= P[-0.94 < Z < 0] + P[0 < Z < 0.31]$

$= P[0 < Z < 0.94] + P[0 < Z < 0.31]$

$= 0.3264 + 0.1217 = 0.4481$

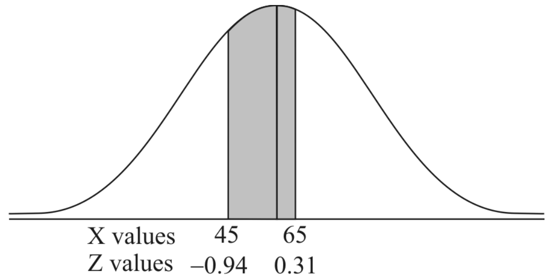


Fig. 14.13: Area Representing Students having Weight between 45 kg and 65 kg

$\therefore$ Number of male students having weight between 45 kg & 65 kg

$$= P[45 < X < 65]$$

$$= 1000 \times 0.4481 = 448$$

b) Let x_1 be the lowest weight amongst 100 heaviest students.

Now, for $X = x_1$, $Z = \dfrac{x_1 - 60}{16} = z_1 \text{(say)}$.

$P[X \geq x_1] = \dfrac{100}{1000} = 0.1$ [See Fig.14.14]

$\Rightarrow P[Z \geq z_1] = 0.1$

$\Rightarrow P\left[0 \leq Z \leq z_1\right] = 0.5 - 0.1 - 0.4.$

$\Rightarrow z_1 = 1.28$ [From Table]

$\Rightarrow x_1 = 60 + 16 \times 1.28 = 60 + 20.48 = 80.48.$

Therefore, the lowest weight of 100 heaviest male students is 80.48 kg.

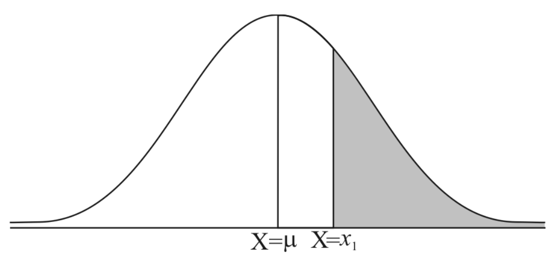


Fig. 14.14: Area Representing the 100 Heaviest Male Students

Example 4: In a normal distribution 10% of the items are over 125 and 35% are under 60. Find the mean and standard deviation of the distribution.

Solution:

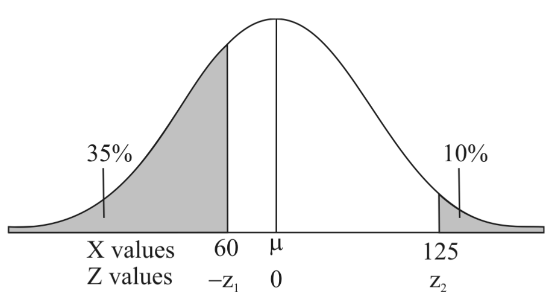


Fig. 14.15: Area Representing the Items under 60 and over 125

Let $X \sim N(\mu, \sigma^2)$, where μ and σ^2 are unknown and are to be obtained.

Here we are given

P[X > 125] = 0.1 and P[X < 60] = 0.35. [See Fig. 14.15]

We know that if X ~ N(μ, σ²), then $Z = \frac{X-\mu}{\sigma}$.

For X = 60, $Z = \frac{60-\mu}{\sigma} = -z_1 \text{(say)}$ ... (1) $\left[\begin{matrix}-\text{ve sign is taken because} \\ P[Z<0] = P[Z>0] = 0.5\end{matrix}\right]$

For X = 125, $Z = \frac{125-\mu}{\sigma} = z_2 \text{(say)}$ ...(2)

Now $P\left[X<60\right] = P\left[Z<-z_1\right] = 0.35$

$\Rightarrow P[Z>z_1] = 0.35$ [By symmetry of normal curve]

$\Rightarrow 0.5 - P[0<Z<z_1] = 0.35$

$\Rightarrow P[0<Z<z_1] = 0.15$

$\Rightarrow z_1 = 0.39$ $\left[\begin{matrix}\text{From the table areas} \\ \text{under normal curve}\end{matrix}\right]$

and $P\left[X>125\right] = P\left[Z>z_2\right] = 0.10$

$\Rightarrow 0.5 - P[0<Z<z_2] = 0.10$

$\Rightarrow P[0<Z<z_2] = 0.40$

$\Rightarrow z_2 = 1.28$ $\left[\text{From the table}\right]$

Putting the values of z_1 and z_2 in Equations (1) and (2), we get

$$\frac{60-\mu}{\sigma} = -0.39 \quad \dots (3)$$

$$\frac{125-\mu}{\sigma} = 1.28 \quad \dots (4)$$

(4) – (3) gives

$$\frac{125-\mu-60+\mu}{\sigma} = 1.28 + 0.39$$

$$\frac{65}{\sigma} = 1.67 \Rightarrow \sigma = \frac{65}{1.67} = 38.92$$

From Eq. (4), $\mu = 125 - 1.28\sigma \Rightarrow \mu = 125 - 1.28 \times 38.92 = 75.18$

Hence $\mu = \text{mean} = 75.18; \sigma = \text{S.D.} = 38.92$

Example 5: Find the quartile deviation of the normal distribution having mean μ and variance σ^2.

Solution: Let X ~ N(μ, σ²). Let Q_1 and Q_3 are the first and third quartiles. Now as Q_1, Q_2 and Q_3 divide the distribution into four equal parts, therefore, areas under the normal curve to the left of Q_1, between Q_1 and Q_2 (Median), between Q_2 and Q_3 and to the right of Q_3 all are equal to 25 percent of the total area. This has been shown in Fig. 14.16.

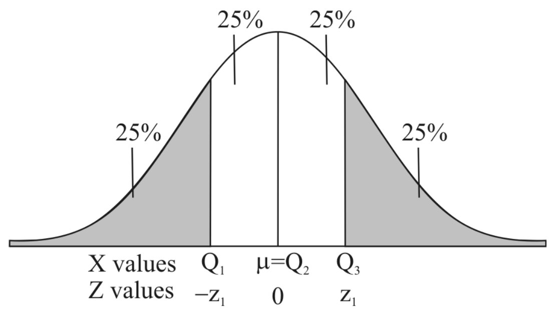


Fig. 14.16: Area to the Left of X = Q_1 and to the Right of X = Q_3

i.e. here, we have

$P[X < Q_1] = 0.25$, $P[Q_1 < X < \mu] = 0.25$, $P[\mu < X < Q_3] = 0.25$ and

$P[X > Q_3] = 0.25$ [See Fig. 14.16]

Now, when $X = Q_1,\ Z = \dfrac{Q_1 - \mu}{\sigma} = -z_1$, (say)

∵ value of Z corresponds to Q_1 which lies to the left of mean which is zero for Z and hence the value to the left of it is negative. Thus, a negative value of Z has been taken here.

$\Rightarrow Q_1 - \mu = -\sigma z_1 \quad \Rightarrow Q_1 = \mu - \sigma z_1$

and when

$X = Q_3,\ Z = \dfrac{Q_3 - \mu}{\sigma} = z_1$

Due to symmetry of normal curve, the values of Z corresponding to Q_1 and Q_3 are equal in magnitude because they are equidistant from mean.

$\Rightarrow Q_3 - \mu = \sigma z_1 \qquad \Rightarrow Q_3 = \mu + \sigma z_1$

Now, as $P[\mu < X < Q_3] = 0.25$, therefore,

$P[0 < Z < z_1] = 0.25$

$\Rightarrow z_1 = 0.67$ [From normal tables]

Now, Q.D. $= \dfrac{Q_3 - Q_1}{2} = \dfrac{(\mu + \sigma z_1) - (\mu - \sigma z_1)}{2} = \sigma z_1 = \sigma(0.67)$ i.e. $\dfrac{2}{3}\sigma$ (approx).

Now, we are sure that you can try the following exercises:

E1) If X ~ N(150, 9) and Z is a S.N.V. i.e $Z = \dfrac{X - \mu}{\sigma}$ then find Z scores corresponding to the following values of X

(i) X = 165 (ii) X = 120

E2) Suppose X ~ N (25, 4) then find

(i) $P[X < 22]$, (ii) $P[X > 23]$, (iii) $P[|X - 24| < 3]$, and (iv) $P[|X - 21| > 2]$

E3) Suppose $X \sim N(30, 16)$ then find α in each case

(i) $P[X > \alpha] = 0.2492$

(ii) $P[X < \alpha] = 0.0496$

E4) Let the random variable X denote the chest measurements (in cm) of 2000 boys, where $X \sim N(85, 36)$.

a) Then find the number of boys having chests measurement
   i) less than or equal to 87 cm,
   ii) between 86 cm and 90 cm,
   iii) more than 80 cm.

b) What is the lowest value of the chest measurement among the 100 boys having the largest chest measurements?

E5) In a particular branch of a bank, it is noted that the duration/waiting time of the customers for being served by the teller is normally distributed with mean 5.5 minutes and standard deviation 0.6 minutes. Find the probability that a customer has to wait

a) between 4.2 and 4.5 minutes, (b) for less than 5.2 minutes, and (c) more than 6.8 minutes

E6) Suppose that temperature of a particular city in the month of March is normally distributed with mean $24^\circ C$ and standard deviation $6^\circ C$. Find the probability that temperature of the city on a day of the month of March is

(a) less than $20^\circ C$ (b) more than $26^\circ C$ (c) between $23^\circ C$ and $27^\circ C$

14.3 FITTING OF NORMAL CURVE USING AREA PROPERTY

To fit a normal curve to the observed data we first find the mean and variance from the given data. Mean and variance so obtained are μ and σ respectively. Substituting these values of μ and σ^2 in the probability function

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, we get the normal curve fitted to the given data.

Now, the expected frequencies can be computed using either of the following two methods:

1. Area method
2. Method of ordinates

But, here we only deal with the area method. Process of finding the expected frequencies by area method is described in the following steps:

(i) Write the lower limits of each of the given class intervals.

(ii) Find the standard normal variate $Z = \frac{X - \mu}{\sigma}$ corresponding to each lower limit. Suppose the values of the standard normal variate are obtained as z_1, z_2, z_3, …

(iii) Find $P[Z \le z_1]$, $P[Z \le z_2]$, $P[Z \le z_3]$,…i.e. the areas under the normal curve to the left of ordinate at each value of Z obtained in step (ii). Using table given in the Appendix at the end of the unit $Z = z_i$ may be to the right or left of $Z = 0$.

If $Z = z_i$ is to the right of $Z = 0$ as shown in the following figure:

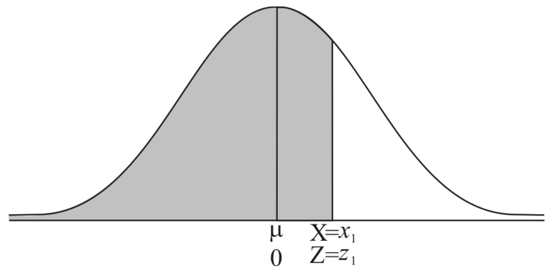


Fig. 14.17: Area to the Left of $Z = Z_i$, when Z_i is to the Right of Z = 0

Then, $P[Z \le z_i]$ is obtained as

$P[Z \le z_i] = 0.5 + P[0 \le Z \le z_i]$

But, if $Z = z_i$ is to the left of $Z = 0$ (this is the case when z_i is negative) as shown in the following figure:

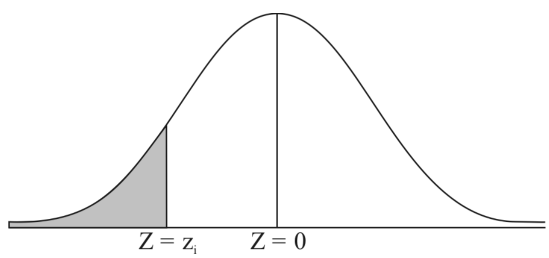


Fig. 14.18: Area to the Left of Z = Z_i, when Z_i is to the left of Z = 0

Then

$P[Z \le z_i] = 0.5 - P[z_i \le Z \le 0]$

$= 0.5 - P[0 \le Z \le -z_i]$ [Due to symmetry]

e.g. $z_i = -2$ (say),

Then $P[Z \le -2] = 0.5 - P[-2 \le Z \le 0]$

$= 0.5 - P[0 \le Z \le -(-2)]$

$= 0.5 - P[0 \le Z \le 2]$

(iv) Obtain the areas for the successive class intervals on subtracting the area corresponding to every lower limit from the area corresponding to the succeeding lower limit.

e.g. suppose 10, 20, 30 are three successive lower limits.

Then areas corresponding to these limits are

$P[X \le 10]$, $P[X \le 20]$, $P[X \le 30]$ respectively.

Now the difference $P[X \le 30] - P[X \le 20]$ gives the area corresponding to the interval 20-30.

(v) Finally, multiply the differences obtained in step (iv) i.e. areas corresponding to the intervals by N (the sum of the observed frequencies), we get the expected frequencies.

Above procedure is explained through the following example.

Example 6: Fit a normal curve by area method to the following data and find the expected frequencies.

X	f
0-10	3
10-20	5
20-30	8
30-40	3
40-50	1

Solution: First we are to find the mean and variance of the given frequency distribution. This you can obtain yourself as you did in Unit 2 of MST-002 and at many other stages. So, this is left an exercise for you.

You will get the mean and variance as

$\mu = 22$ and $\sigma^2 = 111$ respectively

$\Rightarrow \sigma = 10.54$

Hence, the equation of the normal curve is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{1}{(10.54)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-22}{10.54}\right)^2}, -\infty < x < \infty$$

Expected frequencies are computed as follows:

Class Interval	Lower Limit	Standard Normal Variate $Z = \frac{X-\mu}{\sigma}$ $= \frac{X-22}{10.54}$	Area under normal curve to the left of Z $P[X < x]$ $= P[Z \le z]$	Difference between successive areas	Expected frequencies $=20 \times \text{col. V}$
Below 0	$-\infty$	$-\infty$	$P[Z < -\infty]$ $= 0$	0.0183 – 0 = 0.0183	0.366 □ 0
0-10	0	–2.09	$P[Z \le -2.09]$ $= 0.0183$	0.1271 – 0.0183 = 0.1088	2.176 □ 2
10-20	10	–1.14	$P[z \le -1.14]$ $= 0.1271$	0.4241 – 0.1271 = 0.2970	5.94 □ 6
20-30	20	–0.19	$P[Z \le -0.19]$ $= 0.4241$	0.7764 – 0.4241 = 0.3523	7.05 □ 7
30-40	30	0.76	$P[Z \le 0.76]$ $= 0.9564$	0.9564 – 0.7764 = 0.1800	3.6 □ 4
40-50	40	1.71	$P[Z \le 1.71]$ $= 0.9564$	0.9961 – 0.9564 = 0.0397	0.79 □ 1
50 and above	50	2.66	$P[Z \le 2.66]$ $= 0.9961$	–	–

The areas under the normal curve shown in the fourth column of the above tables are obtained as follows:

$P[Z < -\infty] = 0$ $\quad \left[\begin{array}{l} \because \text{there is no value} \\ \text{to the left of } -\infty \end{array}\right]$

$P[Z \le -2.09] = 0.5 - P[-2.09 \le Z \le 0]$ [See Fig. 14.19]

$= 0.5 - P[0 \le Z \le 2.09]$ [Due to symmetry]

$= 0.5 - 0.4817$ $\quad \left[\begin{array}{l} \text{From table given at} \\ \text{the end of the unit} \end{array}\right]$

$= 0.0183$

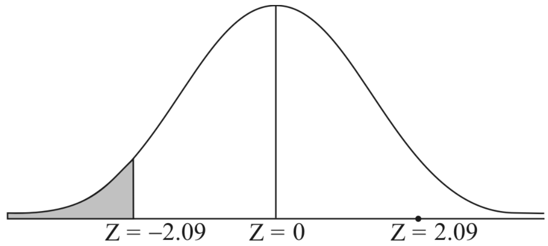


Fig. 14.19: Area to the Left of Z = –2.09

Similarly,

$P[Z \le -1.14] = 0.5 - 0.3729 = 0.1271$

$P[Z \le -0.19] = 0.5 - 0.0759 = 0.4241$

Now, $P[Z \le 0.76] = 0.5 + P[0 \le Z \le 0.76]$ [See Fig. 14.20]

$= 0.5 + 0.2764$

$= 0.7764$

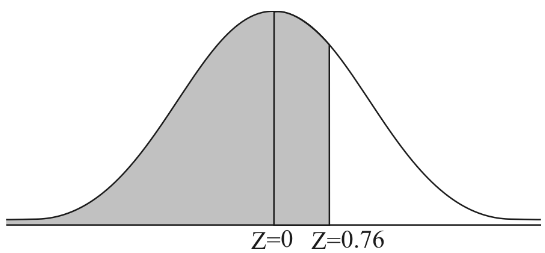


Fig. 14.20: Area to the Left of Z = 0.76

Similarly

$P[Z \le 1.71] = 0.5 + 0.4564 = 0.9564$

$P[Z \le 2.66] = 0.5 = 0.4961 = 0.9961$

You can now try the following exercises:

E7) Fit a normal curve to the following distribution and find the expected frequencies by area method.

X	60– 65	65-70	70-75	75-80	80-85
	5	8	12	8	7

E8) The following table gives the frequencies of occurrence of a variate X between certain limits. The distribution is normal. Find the mean and S.D. of X.

X	Less than 40	40-50	50 and more
f	30	33	37

14.4 SUMMARY

The main points covered in this unit are:

1) **Area property and its various applications** has been discussed in detail.
2) **Quartile deviation** has also been obtained using the **area property** in an example.
3) **Fitting of normal distribution** using area property and computation of **expected frequencies** using area method have been explained.

14.5 SOLUTIONS/ANSWERS

E1) We are given X ~ N(150, 9)

$\therefore$ in usual notations, we have

$\mu = 150,\ \sigma^2 = 9 \quad \Rightarrow \sigma = 3$

Now, $Z = \frac{X-\mu}{\sigma} = \frac{X-150}{3}$

(i) When X = 165, $Z = \frac{165-150}{3} = \frac{15}{3} = 5$

(ii) When X = 120, $Z = \frac{120-150}{3} = \frac{-30}{3} = -10$

E2) Here X ~ N(25, 4)

$\therefore$ in usual notations, we have

$\text{Mean} = \mu = 25, \quad \text{variance} = \sigma^2 = 4 \Rightarrow \sigma = 2$

If Z is the S.N.V then $Z = \frac{X-\mu}{\sigma} = \frac{X-25}{2}$

i) $X = 22, \quad Z = \frac{22-25}{2} = \frac{-3}{2} = -1.5$

$P[X < 22] = P[Z < -1.5]$ [See Fig. 14.21]

$= P[Z > 1.5]$ $\left[\because \text{due to symmetry of normal curve}\right]$

$= 0.5 - P[0 < Z < 1.5]$

$= 0.5 - 0.4332$ $\left[\text{Using table area under normal curve}\right]$

$= 0.0668$

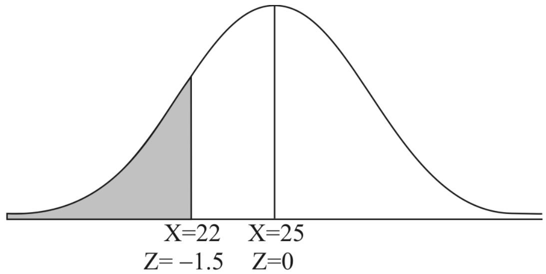


Fig. 14.21: Area to the Left of X = 22

ii) $X = 23, \quad Z = \dfrac{23-25}{2} = \dfrac{-2}{2} = -1$

$P[X > 23] = P[Z > -1]$ [See Fig.14.22]

$= P[Z < 1]$ $\left[\because \text{ due to symmetry of normal curve}\right]$

$= 0.5 + P[0 < Z < 1]$

$= 0.5 + 0.3413$ $\left[\text{Using table area under normal curve}\right]$

$= 0.8413$

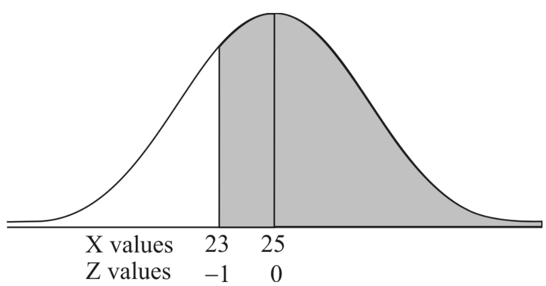


Fig. 14.22: Area to the Right of X = 23

iii) $P[|X-24| < 3] = P[-3 < X - 24 < 3]$ $\left[\because |x-a| < b \Rightarrow -b < x-a < b\right]$

$= P[-3+24 < X < 3+24)$

$= P[21 < X < 27]$

$X = 21, \quad Z = \dfrac{21-25}{2} = \dfrac{-4}{2} = -2$

$X = 27, \quad Z = \dfrac{27-25}{2} = \dfrac{2}{2} = 1$

$\therefore P[|X-24| < 3] = P[21 < X < 27]$ [See Fig.14.23]

$= P[-2 < Z < 1]$

$$= P[-2 < Z < 0] + P[0 < Z < 1]$$

$$= P[0 < Z < 2] + P[0 < Z < 1]$$

$$= 0.4772 - 0.3413 = 0.1359$$

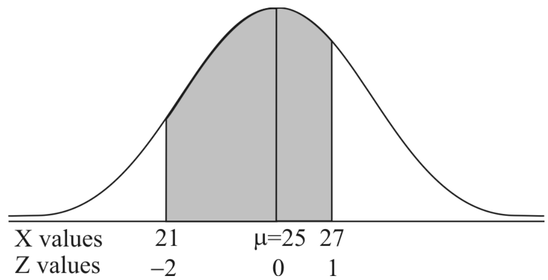


Fig. 14.23: Area between X = 21 and X = 27

iv) $P[|X-21|>2] = P[X-21>2 \text{ or } -(X-21)>2]$

$$\left[\begin{array}{l}\because |x-a|>b \Rightarrow \pm(x-a)>b \\ \Rightarrow x-a>b \text{ or } -(x-a)>b\end{array}\right]$$

$$= P[X>23 \text{ or } -X>2-21]$$

$$= P[X>23 \text{ or } X<19] \qquad \left[\begin{array}{l}\because -y>-a \\ \Rightarrow y<a\end{array}\right]$$

For X=19, $Z = \dfrac{19-25}{2} = \dfrac{-6}{2} = -3$

For X=23, $Z = \dfrac{23-25}{2} = \dfrac{-2}{2} = -1$

$\therefore P[|X-21|>2] = P[X>23 \text{ or } X<19]$ $\qquad$ [See Fig 14.24]

$$= P[Z>-1 \text{ or } Z<-3]$$

$$= P[Z>-1] + P[Z<-3] \qquad \left[\begin{array}{l}\text{By addition theorem for} \\ \text{mutually exclusive events}\end{array}\right]$$

$$= 1 - P[-3<Z<-1]$$

$$= 1 - P[1<Z<3]$$

$$= 1 - [P[0<Z<3] - P[0<Z<1]]$$

$$= 1 - [0.4987 - 0.3413] \qquad [\text{From table}]$$

$$= 1 - 0.1574 = 0.8426.$$

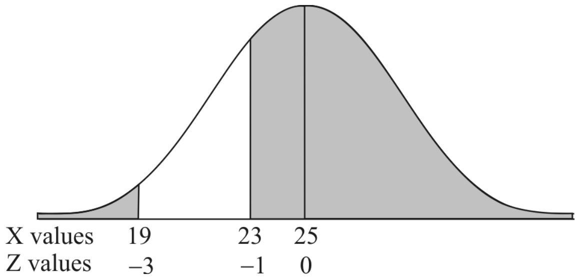


Fig. 14.24: Area between X = 19 and X = 23

E3) Here X ~ N(30, 16)

$\therefore$ in usual notations, we have

Mean = $\mu = 30$, variance $= \sigma^2 = 16 \Rightarrow \sigma = 4$

If Z is S.N.V then $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-30}{4}$

i) $X = \alpha$, $\quad Z = \dfrac{\alpha - 30}{4} = z_1(\text{say}) \qquad \ldots(1)$

Now $P[X > \alpha] = 0.2492$ [See Fig.14.25]

$\Rightarrow P[Z > z_1] = 0.2492 \Rightarrow 0.5 - P\left[0 < Z < z_1\right] = 0.2492$

$\Rightarrow P\left[0 < Z < z_1\right] = 0.2508$

$\Rightarrow z_1 = 0.67$ [From the table]

Putting $z_1 = 0.67$ in (1), we get

$\dfrac{\alpha - 30}{4} = 0.67$

$\alpha - 30 = 2.68$

$\alpha = 30 + 2.68 = 32.68$

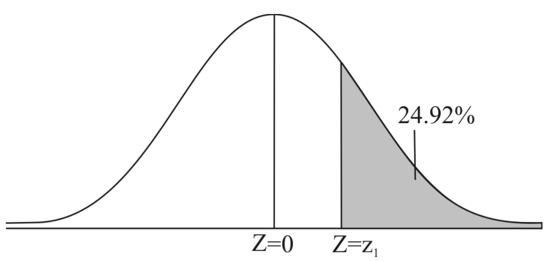


Fig. 14.25: z_1 Corresponding to 24.92 % Area to its Right

ii) For $X = \alpha$, $\quad Z = \dfrac{\alpha - 30}{4} = -z_2(\text{say}) \quad \ldots (2)$

Now $P[X < \alpha] = 0.0496$

$\Rightarrow P[Z < -z_2] = 0.0496$ [See Fig.14.26]

$\Rightarrow P[Z > z_2] = 0.0496$ [Due to symmetry]

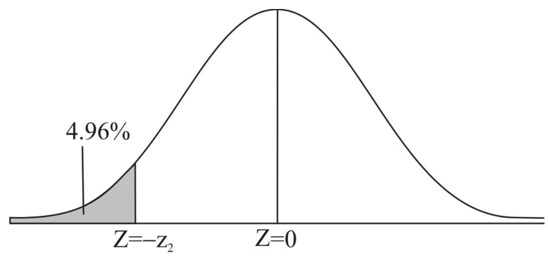


Fig. 14.26: z_2 Corresponding to 4.96 % Area to its Right

$$\Rightarrow 0.5 - P[0 < Z < z_2] = 0.0496$$

$$\Rightarrow P[0 < Z < z_2] = 0.5 - 0.0496 = 0.4504$$

$$\Rightarrow z_2 = 1.65 \qquad \text{[From the table]}$$

Putting $\Rightarrow z_2 = 1.65$ in (2), we get

$$\frac{\alpha - 30}{4} = -1.65$$

$$\Rightarrow \alpha - 30 = -1.65 \times 4$$

$$\Rightarrow \alpha - 30 = -6.60$$

$$\Rightarrow \alpha = 30 - 6.6 = 23.4$$

E4) We are given X ~ N(85, 36), N = 2000

i.e. $\mu = 85\text{cm}, \sigma^2 = 36\text{cm}, N = 2000$

If X ~ N(μ, σ^2) and $Z = \dfrac{x - \mu}{\sigma}$ then we know that Z ~ N(0, 1)

a) i) For X = 87, $Z = \dfrac{87 - 85}{6} = \dfrac{2}{6} \approx 0.33$

Now P[X < 87] = P [Z < 0.33] [See Fig. 14.27]

= 0.5 + P [0 < Z < 0.33]

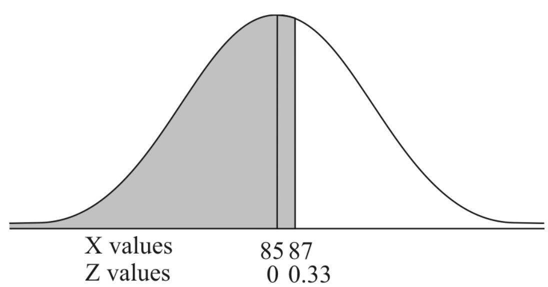


Fig. 14.27: Area to the Left of X = 87 or Z = 0.33

= 0.5 + P [0 < Z < 0.33]

$$= 0.5 + 0.1293 \qquad \begin{bmatrix}\text{From the table of areas} \\ \text{under normal curve}\end{bmatrix}$$

= 0.6293

Therefore, number of boys having chests measurement ≤ 87

$$= N.P[X \le 87]$$

$$= 2000 \times 0.6293 = 1259$$

ii) For X = 86, $Z = \dfrac{86-85}{6} = \dfrac{1}{6} \approx 0.17$

For X = 90, $Z = \dfrac{90-85}{6} = \dfrac{5}{6} \approx 0.83$

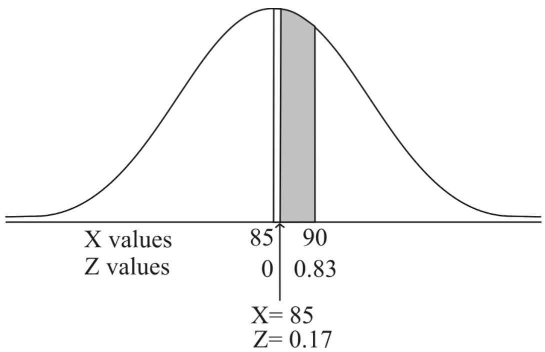


Fig. 14.28: Area between X= 86 and X= 90

$$\therefore \; P[86 < X < 90] = P[0.17 < Z < 0.83] \quad [\text{See Fig.14.28}]$$

$$= P[0 < Z < 0.83] - P[0 < Z < 0.17]$$

$$= 0.2967 - 0.0675 \qquad \left[\begin{array}{l}\text{From the table of areas}\\ \text{under normal curve}\end{array}\right]$$

$$= 0.2292$$

$\therefore$ number of boys having chests measurement between 86 cm and 90 cm

$$= N.\ P\ [86 < x < 90]$$

$$= 2000 \times 0.2292 = 458$$

iii) For X = 80, $Z = \dfrac{80-85}{6} = \dfrac{-5}{6} \approx -0.83$

P [X > 80] = P [Z > – 0.83] [See Fig. 14.29]

$$= P\ [Z < 0.83]$$

$$= 0.5 + P\ [0 < Z < 0.83]$$

$$= 0.5 + 0.2967 \qquad \left[\begin{array}{l}\text{From the table of areas}\\ \text{under normal curve}\end{array}\right]$$

$$= 0.7967$$

$\therefore$ number of boys having chest measurement more than 80 cm

$$= N.P[X > 80]$$

$$= 2000 \times 0.7967$$

$$= 1593$$

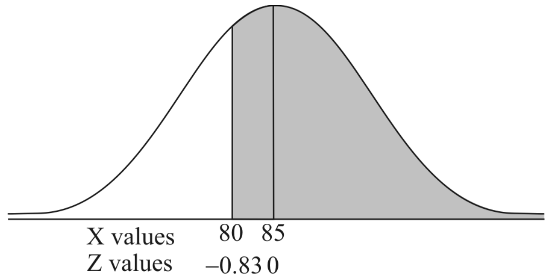


Fig. 14.29: Area to the Right of X = 80 or Z = − 0.83

b) Let x_1 be the lowest chest measurement amongst 100 boys having the largest chest measurements.

Now, for $X = x_1$, $Z = \dfrac{x_1 - 85}{6} = z_1 \text{(say)}$.

$$P[X \geq x_1] = \frac{100}{2000} = 0.05$$

$\Rightarrow P\left[Z \geq z_1\right] = 0.05$ $\qquad\qquad$ [See Fig.14.30]

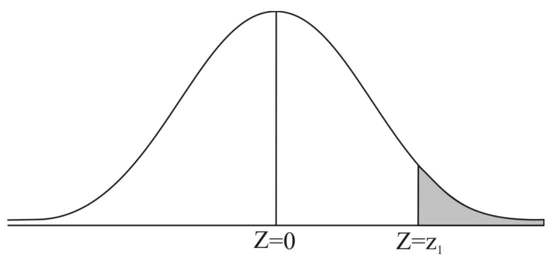


Fig. 14.30: Area Representing the 100 Boys having Largest Chest Measurements

$\Rightarrow P[0 \leq Z \leq z_1] = 0.5 - 0.05 = 0.45$

$\Rightarrow z_1 = 1.64$ $\qquad\qquad$ [From Table]

$\Rightarrow x_1 = 85 + 6 \times 1.64 = 85 + 9.84 = 94.84.$

Therefore, the lowest value of the chest measurement among the 100 boys having the largest chest measurement is 94.84 cm.

E5) We are given

$\mu = 5.5$ minutes, $\sigma = 0.6$ minutes

If $X \sim N(\mu, \sigma^2)$ and $Z = \dfrac{X - \mu}{\sigma}$ then we know that $Z \sim N(0, 1)$

a) For X = 4.2, $Z = \dfrac{4.2 - 5.5}{0.6} = \dfrac{-1.3}{0.6} = \dfrac{-13}{6} \approx -2.17$

For X = 4.5, $Z = \dfrac{4.5 - 5.5}{0.6} = \dfrac{-1.0}{0.6} = \dfrac{-10}{6} = \dfrac{-5}{3} \approx -1.67$

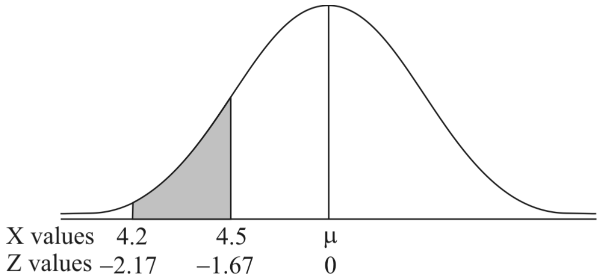


Fig. 14.31: Area Representing Probability of Waiting Time between 4.2 and 4.5 Minutes

$$P[4.2 < x < 4.5] = P[-2.17 < Z < -1.67] \quad [\text{See Fig.14.31}]$$

$$= P\,[1.67 < Z < 2.17]$$

$$= P\,[0 < Z < 2.17] - P\,[0 < Z < 1.67]$$

$$= 0.4850 - 0.4525$$

$$= 0.0325$$

Therefore, probability that customer has to wait between 4.2 min and 4.5 min = 0.0325

b) For $X = 5.2,\ Z = \dfrac{5.2-5.5}{0.6} = \dfrac{0.3}{0.6} = \dfrac{-3}{6} = \dfrac{-1}{2} = -0.5$

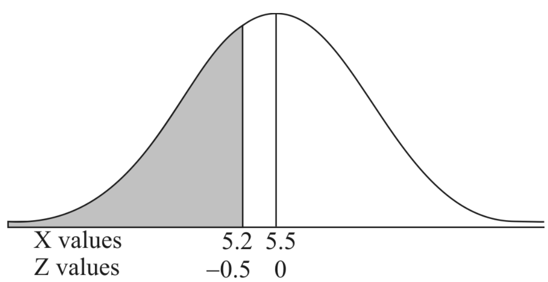


Fig. 14.32: Area Representing Probability of Waiting Time Less than 5.2 Minutes

$$P[X < 5.2] = P\,[Z < -0.5] \qquad [\text{See Fig 14.32}]$$

$$= P\,[Z > 0.5]$$

$$= 0.5 - P\,[0 < Z < 0.5]$$

$$= 0.5 - 0.1915 \qquad \begin{bmatrix}\text{From the table of areas} \\ \text{under normal curve}\end{bmatrix}$$

$$= 0.3085$$

Therefore, probability that customer has to work for less than 5.2 min

= 0.3085

c) For $X = 6.8, Z = \dfrac{6.8-5.5}{0.6} = \dfrac{1.3}{0.6} = \dfrac{13}{6} \approx 2.17$

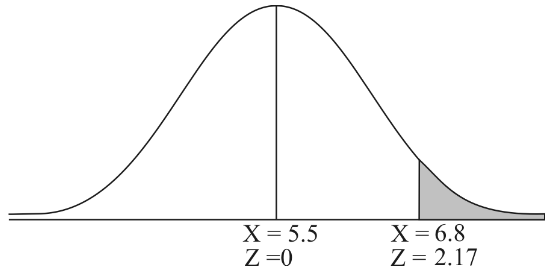


Fig. 14.33: Area Representing Probability of Waiting Time Greater than 6.8 Minutes

$$P[X > 6.8] = P\,[Z > 2.17] \qquad \text{[See Fig 14.33]}$$
$$= 0.5 - P\,[0 < Z < 2.17]$$
$$= 0.5 - 0.4850 = 0.0150$$

Therefore, probability that customer has to wait for more than 6.8 min = 0.0150

E6) Let the random variable X denotes the temperature of the city in the month of March. Then we are given

$X \sim N(\mu, \sigma^2),$ where $\mu = 24^\circ C, \sigma = 6^\circ C$

We know that if $X \sim N(\mu, \sigma^2),$ and $Z = \dfrac{X-\mu}{\sigma}$ then Z ~ N (0, 1)

a) For $X = 20, Z = \dfrac{20-24}{6} = \dfrac{-4}{6} = \dfrac{-2}{3} \approx -0.67$

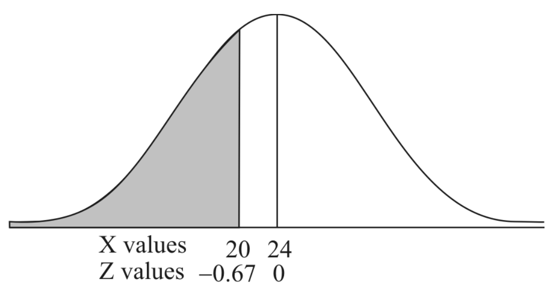


Fig. 14.34: Area Representing Probability of Temperature Less than 20 $^\circ$C

$$P[X < 20] = P[Z < -0.67] \qquad \text{[See Fig. 14.34]}$$
$$= P\,[Z > 0.67]$$
$$= 0.5 - P[0 < Z < 0.67]$$
$$= 0.5 - 0.2486 = 0.2514$$

Therefore, probability that temperature of the city is less than 20 $^\circ$C is 0.2514

b) For X= 26, $Z = \dfrac{26-24}{6} = \dfrac{2}{6} = \dfrac{1}{3} \approx 0.33$

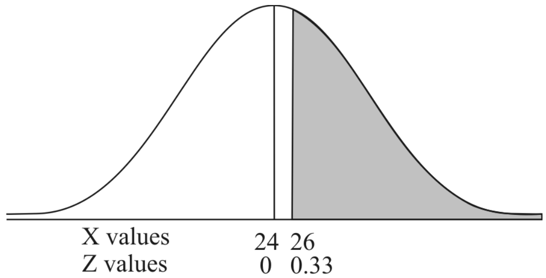


Fig. 14.35: Area Representing Probability of Temperature Greater than 26 $^{\circ}$C

Since, P[X > 26] = P [Z > 0.33] [See Fig. 14.35]

$= 0.5 - P[0 < Z < 0.33]$

$= 0.5 - 0.1293 \begin{bmatrix} \text{From the table of areas} \\ \text{under normal curve} \end{bmatrix}$

$= 0.3707$

Therefore, probability that temperature of the city is more than $26^{\circ}C$ is 0.3707

c) For X= 23, $Z = \dfrac{23-24}{6} = \dfrac{-1}{6} \approx -0.17$

For X= 27, $Z = \dfrac{27-24}{6} = \dfrac{3}{6} = \dfrac{1}{2} = 0.5$

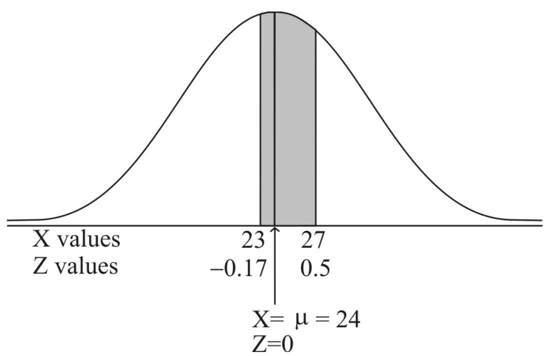


Fig. 14.36: Area Representing Probability of Temperature between 23 $^{\circ}$C and 27 $^{\circ}$C

P[23 < X < 27] = P[–0.17 < Z < 0.5] [See Fig. 14.36]

$= P[-0.17 < Z < 0] + P[0 < Z < 0.5]$

$= P[0 < Z < 0.17] + P[0 < Z < 0.5]$

$= 0.0675 + 0.1915 \quad \begin{bmatrix} \text{From the table of areas} \\ \text{under Normal Curve} \end{bmatrix}$

$= 0.2590$

Therefore, probability that temperature of the city is between $23^{\circ}C$ and $27^{\circ}C$ is 0.2590

E7) Mean $(\mu) = 73$, variance$(\sigma^2) = 39.75$

and hence S.D. $(\sigma) = 6.3$

$\therefore$ The equation of the normal curve fitted to the given data is

$$f(x) = \frac{1}{(6.3)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-73}{6.3}\right)^2}, \qquad -\infty < x < \infty$$

Using area method,

The expected frequencies are obtained as follows:

Class interval	Lower limit X	$Z = \frac{X-\mu}{\sigma} = \frac{X-73}{6.3}$	Area under normal curve to the left of z	Difference between successive areas	Expected frequency $40 \times$ col. V
Below 60	$-\infty$	$-\infty$	0	0.0197 – 0 = 0.0197	$0.8 \approx 1$
60 – 65	60	–2.06	0.5 – 0.4803 = 0.0197	0.1020 – 0.0197 = 0.0823	$3.3 \approx 3$
65 – 70	65	–1.27	0.5 – 0.3980 = 0.1020	0.3156 – 0.1020 = 0.2136	$8.5 \approx 9$
70 – 75	70	–0.48	0.5 – 0.1844 = 0.3156	0.6255 – 0.3156 = 0.3099	$12.4 \approx 12$
75 – 80	75	+ 0.32	0.5 + 0.1255 = 0.6255	0.8655 – 0.6255 = 0.2400	$9.6 \approx 10$
80 – 85	80	1.11	0.5 + 0.3655 = 0.8655	0.9713 – 0.8655 = 1.1058	$4.2 \approx 4$
85 and above	85	1.90	0.5 + 0.4713 = 0.9713		

E8) $P[X < 40] = \frac{30}{100} = 0.3,$

$P[40 < X < 50] = \frac{33}{100} = 0.33,$ and

$$P[X > 50] = \frac{37}{100} = 0.37,$$

Now, Let $X \sim N(\mu, \sigma^2)$,

$\therefore$ Standard normal variate is

$$Z = \frac{X - \mu}{\sigma}$$

When X = 40, $Z = \frac{40 - \mu}{\sigma} = -z_1$, (say) $\left[\begin{array}{l}\text{It is taken as} - \text{ve as area to}\\ \text{the left of this value is } 30\%\\ \text{as probability is } 0.3\end{array}\right]$

When X = 50, $Z = \frac{50 - \mu}{\sigma} = z_2$ (say) $\left[\begin{array}{l}\text{It is taken as +ve as}\\ \text{area to the right of this}\\ \text{value is given as } 37\%\end{array}\right]$

Now,

$P[X < 40] = P[Z < -z_1] = 0.3$

$\Rightarrow 0.5 - P[-z_1 < Z < 0] = 0.3$ [See Fig 14.37]

$\Rightarrow 0.5 - P[0 < Z < z_1] = 0.3$ [Due to symmetry]

$\Rightarrow P[0 < Z < z_1] = 0.2$

From table at the end of this unit,

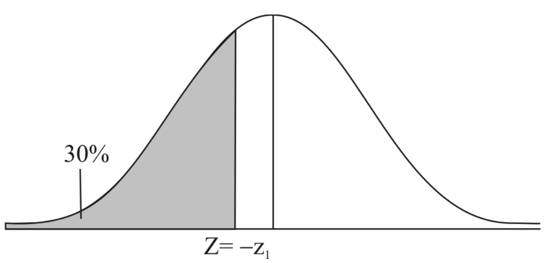


Fig. 14.37: -z_1 Corresponding to the 30% Area to its Left

The value of Z corresponding to probability/area is

$z_1 = 0.525$ $\left[\begin{array}{l}\text{As values of } Z \text{ are } 0.52 \text{ and } 0.53\\ \text{corresponding to the proabability}\\ 0.1985\end{array}\right]$

$P[X > 50] = 0.37$

$\Rightarrow P[Z > z_2] = 0.37$ [See Fig.14.38]

$\Rightarrow 0.5 - P[0 < Z < z_2] = 0.37$

$\Rightarrow P[0 < Z < z_2] = 0.13$

$z_2 = 0.33$(approx.) [From the table]

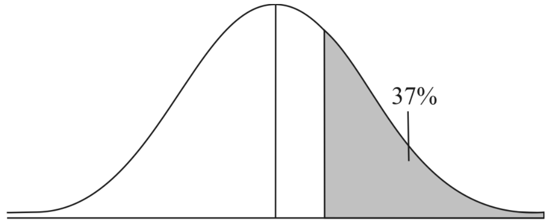


Fig. 14.38: z_2 Corresponding to the 37% Area to its Right

$$\therefore \frac{40-\mu}{\sigma} = -0.525 \text{ and } \frac{50-\mu}{\sigma} = 0.33$$

$$\Rightarrow 40-\mu = 0.525\sigma \text{ and } 50-\mu = 0.33\sigma$$

Solving these equations for μ and σ, we have

$\sigma = 11.7$ and $\mu = 46.14$

APPENDIX

AREAS UNDER NORMAL CURVE

The standard normal probability curve is given by

$$\phi(z)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right), -\infty < z < \infty$$

The following table gives probability corresponding to the shaded area as shown in the following figure i.e. $P[0 < Z < z]$ for different values of z

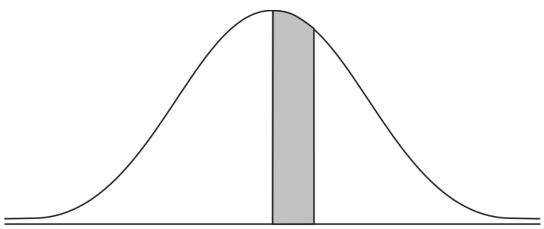


	TABLE OF AREAS									
↓ z →	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0759
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2005	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3655	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3820
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	4279	.4292	.4306	.4319

1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4678	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4959	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.1960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4879	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4493	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

UNIT 15 CONTINUOUS UNIFORM AND EXPONENTIAL DISTRIBUTIONS

Structure

15.1 INTRODUCTION

In Units 13 and 14, you have studied normal distribution with its various properties and applications. Continuing our study on continuous distributions, we, in this unit, discuss continuous uniform and exponential distributions. It may be seen that discrete uniform and geometric distributions studied in Unit 11 and Unit 12 are the discrete analogs of continuous uniform and exponential distributions. Like geometric distribution, exponential distribution also has the memoryless property. You have also studied that geometric distribution is the only discrete distribution which has the memoryless property. This feature is also there in exponential distribution and it is the only continuous distribution having the memoryless property.

The present unit discusses continuous uniform distribution in Sec. 15.2 and exponential distribution in Sec. 15.3.

Objectives

After studing the unit, you would be able to:

- define continuous uniform and exponential distributions;
- state the properties of these distributions;
- explain the memoryless property of exponential distribution; and
- solve various problems on the situations related to these distributions.

15.2 CONTINUOUS UNIFORM DISTRIBUTION

The uniform (or rectangular) distribution is a very simple distribution. It provides a useful model for a few random phenomena like having random number from the interval [0, 1], then one is thinking of the value of a uniformly distributed random variable over the interval [0, 1].

Definition: A random variable X is said to follow a continuous uniform (rectangular) distribution over an interval (a, b) if its probability density function is given by

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

The distribution is called uniform distribution since it assumes a constant (uniform) value for all x in (a, b). If we draw the graph of y = f(x) over x-axis and between the ordinates x = a and x = b (say), it describes a rectangle as shown in Fig. 15.1

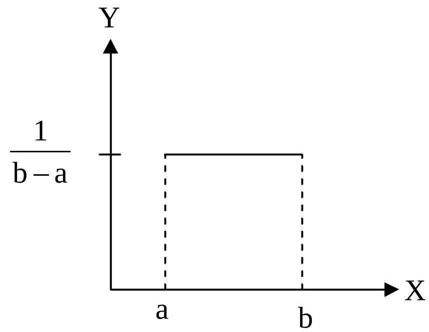


Fig. 15.1: Graph of uniform function

A uniform variate X on the interval (a, b) is written as X ~ U[a, b]

Cumulative Distribution Function

The cumulative distribution function of the uniform random variate over the interval (a, b) is given by:

For $x \le a$, $F(x) = P[X \le x] = \int_{-\infty}^{x} 0dx = 0$

For a < x < b,

$$F(x) = P[X \le x] = \int_{a}^{x} f(x)dx = \int_{a}^{x} \frac{1}{b-a}dx = \frac{1}{b-a}[x]_a^x = \frac{x-a}{b-a}.$$

For $x \ge b$

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)dx$$

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

$$= \int_{-\infty}^{a} (0)dx + \int_{a}^{b} \frac{1}{b-a}dx + \int_{b}^{\infty} (0)dx$$

$$= 0 + \frac{1}{b-a}[x]_a^b + 0 = \frac{b-a}{b-a} = 1.$$

So,

$$F(x) = \begin{cases} 0 & \text{for } x \le a \\ \dfrac{x-a}{b-a} & \text{for } a < x < b \\ 1 & \text{for } x \ge b \end{cases}$$

On plotting its graph, we have

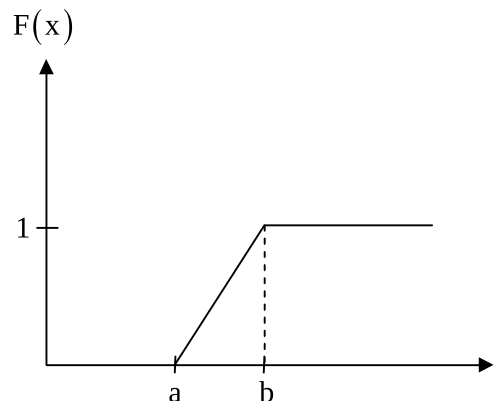


Fig. 15.2: Graph of distribution function

Mean and Variance of Uniform Distribution

Mean = 1st order moment about origin (μ_1')

$$= \int_a^b x.f(x)dx = \int_a^b x.\frac{1}{b-a}dx$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{b-a}\left[\frac{b^2}{2}-\frac{a^2}{2}\right]$$

$$= \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$$

Second order moment about origin (μ_2')

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

$$= \int_a^b x^2.\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{b-a}\left[\frac{b^3}{3}-\frac{a^3}{3}\right]$$

$$= \frac{b^3-a^3}{3(b-a)}$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} \quad \left[\because x^3-y^3=(x-y)(x^2+xy+y^2)\right]$$

$$= \frac{a^2+ab+b^2}{3}$$

$\therefore$ Variance of X = E(X^2) – [E(X)]^2 = $\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2$

$$= \frac{4(a^2+ab+b^2)-3(a+b)^2}{12}$$

$$= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 3b^2 - 6ab}{12}$$

$$= \frac{b^2 + a^2 - 2ab}{12} = \frac{(b-a)^2}{12}.$$

So, Mean = $\frac{a+b}{2}$ and Variance = $\frac{(b-a)^2}{12}$.

Let us now take up some examples on continuous uniform distribution.

Example 1: If X is uniformly distributed with mean 2 and variance 12, find P[X < 3].

Solution: Let X ~ U [a, b]

$\therefore$ probability density function of X is

$$f(x) = \frac{1}{b-a},\ a < x < b.$$

Now as Mean = 2

$$\Rightarrow \frac{a+b}{2} = 2$$

$\Rightarrow$ a + b = 4 … (1)

Variance = 12

$$\Rightarrow \frac{(b-a)^2}{12} = 12$$

$$\Rightarrow (b-a)^2 = 144$$

$\Rightarrow$ b – a = ± 12

$\Rightarrow$ b – a = 12 ... (2) $\left[\begin{array}{l}\because b - a = -12, \text{being negative is} \\ \text{rejected as b should be greater than a} \\ \Rightarrow b - a\ \text{should be positive}\end{array}\right]$

Adding (1) and (2), we have

2b = 16

$\Rightarrow$ b = 8 and hence a = 4 – 8 = – 4

$$\therefore\ f(x) = \frac{1}{b-a} = \frac{1}{8-(-4)} = \frac{1}{12} \text{ for } -4 < x < 8.$$

Thus, the desired probability $= P[X < 3] = \int_{-4}^{3} \frac{1}{12} dx = \frac{1}{12}\int_{-4}^{3} 1 dx = \frac{1}{12}[x]_{-4}^{3}$

$$= \frac{1}{12}[3-(-4)] = \frac{7}{12}.$$

Example 2: Calculate the coefficient of variation for the rectangular distribution in (0, 12).

Solution: Here a = 0, b = 12.

$\therefore$ Mean = $\frac{a+b}{2} = \frac{0+12}{2} = 6,$

Variance = $\frac{(b-a)^2}{12} = \frac{(12-0)^2}{12} = \frac{144}{12} = 12.$

$\Rightarrow$ S.D. = $\sqrt{12}$

Thus, the coefficient of variation

$$= \frac{\text{S.D.}}{\text{Mean}} \times 100 \qquad \text{[Also see Unit 2 of MST-002]}$$

$$= \frac{\sqrt{12}}{6} \times 100 = 57.74\%$$

Example 3: Metro trains are scheduled every 5 minutes at a certain station. A person comes to the station at a random time. Let the random variable X count the number of minutes he/she has to wait for the next train. Assume X has a uniform distribution over the interval (0, 5). Find the probability that he/she has to wait at least 3 minutes for the train.

Solution: As X follows uniform distribution over the interval (0, 5),

$\therefore$ probability density function of X is

$$f(x) = \frac{1}{b-a} = \frac{1}{5-0} = \frac{1}{5}, \; 0 < x < 5$$

Thus, the desired probability

$$P[X \geq 3] = \int_3^5 f(x)\,dx = \int_3^5 \frac{1}{5}\,dx = \frac{1}{5}\int_3^5 (1)\,dx$$

$$= \frac{1}{5}[x]_3^5 = \frac{1}{5}(5-3) = \frac{2}{5} = 0.4$$

Now, you can try the following exercises.

E1) Suppose that X is uniformly distributed over (–a, a). Determine 'a' so that

i) $P[X > 4] = \frac{1}{3}$

ii) $P[X < 1] = \frac{3}{4}$

iii) $P[|X| < 2] = P[|X| > 2]$

E2) A random variable X has a uniform distribution over (–2, 2). Find k for which $P[X > k] = \frac{1}{2}$.

Now, let us discuss exponential distribution in the next section.

15.3 EXPONENTIAL DISTRIBUTION

The exponential distribution finds applications in the situations related to lifetime of an equipment or service time at the counter in a queue. So, the exponential distribution serves as a good model whenever there is a waiting time involved for a specific event to occur e.g. waiting time for a failure to occur in a machine. The exponential distribution is defined as follows:

Definition: A random variable X is said to follow exponential distribution with parameter $\lambda > 0$, if it takes any non-negative real value and its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \ge 0 \\ 0, & \text{elsewhere} \end{cases}$$

Its cumulative distribution function (c.d.f.) is thus given by

$$F(x) = P[X \le x] = \int_0^x f(x)\,dx = \int_0^x \lambda e^{-\lambda x}\,dx$$

$$= \lambda\left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^x = -1\left[e^{-\lambda x}\right]_0^x = -\left(e^{-\lambda x} - e^0\right)$$

$$= -\left(e^{-\lambda x} - 1\right) = 1 - e^{-\lambda x}.$$

$$\text{So, } F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \ge 0 \\ 0, & \text{elsewhere} \end{cases}.$$

Mean and Variance of Exponential Distribution

$$\text{Mean} = E(X) = \int_0^\infty x\, f(x)\,dx = \int_0^\infty x\,\lambda e^{-\lambda x}\,dx$$

$$= \lambda \int_0^\infty x\, e^{-\lambda x}\,dx$$

$$= \lambda\left[\left[(x)\frac{e^{-\lambda x}}{-\lambda}\right]_0^\infty - \int_0^\infty (1)\frac{e^{-\lambda x}}{-\lambda}\,dx\right] \qquad \text{[Integrating by parts]}$$

In case of integration of product of two different types of functions, we do integration by parts i.e. the following formula is applied:

$$\int (\text{First function})(\text{Second fuction})\,dx$$

= (First function as it is) (Integral of second)

$$-\int (\text{Differentiation of first})(\text{Integral of second})\,dx$$

$$\therefore \text{ Mean } = \lambda\left[(0-0)+\frac{1}{\lambda}\left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^\infty\right]$$

$$= \lambda\left[-\frac{1}{\lambda^2}(0-1)\right] = \lambda\left(\frac{1}{\lambda^2}\right) = \frac{1}{\lambda}.$$

Now, $E\left(X^2\right) = \int_0^\infty x^2 f(x)\,dx = \int_0^\infty x^2\left(\lambda e^{-\lambda x}\right)dx$

$$= \lambda\int_0^\infty x^2 e^{-\lambda x}\,dx$$

$$= \lambda\left[\left(x^2\frac{e^{-\lambda x}}{-\lambda}\right)_0^\infty - \int_0^\infty (2x)\frac{e^{-\lambda x}}{-\lambda}\,dx\right] \qquad \text{[Integrating by parts]}$$

$$= \lambda\left[(0-0)+\frac{2}{\lambda}\int_0^\infty x\,e^{-\lambda x}\,dx\right]$$

$$= \frac{2}{\lambda}\lambda\int_0^\infty x\ e^{-\lambda x}\,dx = \frac{2}{\lambda}\int_0^\infty x\left(\lambda e^{-\lambda x}\right)dx$$

$$= \frac{2}{\lambda}E(X)$$

$$= \frac{2}{\lambda}\frac{1}{\lambda} \qquad \text{[E(X) is mean and has already been obtained]}$$

$$= \frac{2}{\lambda^2}$$

Thus, Variance = $E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

So, Mean = $\frac{1}{\lambda}$ and Variance = $\frac{1}{\lambda^2}$.

Remark 1: Variance = $\frac{1}{\lambda^2} = \frac{1}{\lambda.\lambda} = \frac{\text{Mean}}{\lambda} \Rightarrow$ Mean $= \lambda \times$ Variance

So,

Value of λ	Implies
$\lambda < 1$	Mean < Variance
$\lambda = 1$	Mean = Variance
$\lambda > 1$	Mean > Variance

Hence, for exponential distribution,

Mean > or = or < Variance according to whether λ > or = or < 1.

Memoryless Property of Exponential Distribution

Now, let us discuss a very important property of exponential distribution and that is the memoryless (or forgetfulness) property. Like geometric distribution in the family of discrete distributions, exponential distribution is the only distribution in the family of continuous distributions which has memoryless property. The memorless property of exponential distribution is stated as:

If X has an exponential distribution, then for every constant $a \geq 0$, one has

$P[X \leq x + a \mid X \geq a] = P[X \leq x]$ for all x i.e. the conditional probability of waiting up to the time $'x + a'$ given that it exceeds 'a' is same as the probability of waiting up to the time 'x'. To make you understand the above concept clearly let us take the following example: Suppose you purchase a TV set, assuming that its life time follows exponential distribution, for which the expected life time has been told to you 10 years (say). Now, if you use this TV set for say 4 years and then you ask a TV mechanic, without informing him/her that you had purchased it 4 years ago, regarding its expected life time. He/she, if finds the TV set as good as new, will say that its expected life time is 10 years.

So, here, in the above example, 4 years period has been forgotten, in a way, and for this example:

P[life time up to 10 years]

= P[life time up to 14 years | life time exceeds 4 years]

i.e. $P[X \leq 10] = P[X \leq 14 \mid X \geq 4]$

or $P[X \leq 10] = P[X \leq 10 + 4 \mid X \geq 4]$

Here a = 4 and $x = 10$.

Let us now prove the memoryless property of exponential distribution.

Proof: $P[X \leq x + a \mid X \geq a] = \dfrac{[(X \leq x + a) \cap (X \geq a)]}{P[X \geq a]}$ [By conditional probability]

where

$$P[(X \leq x + a) \cap (X \geq a)] = P[a \leq X \leq x + a]$$

$$= \int_a^{x+a} f(x)\,dx = \lambda \int_a^{x+a} e^{-\lambda x}dx$$

$$= \lambda\left[\frac{e^{-\lambda x}}{-\lambda}\right]_a^{x+a} = \lambda\left[\frac{e^{-\lambda(x+a)}}{-\lambda} - \frac{e^{-\lambda a}}{-\lambda}\right]$$

$$= \left[-e^{-\lambda(x+a)} + e^{-\lambda a}\right] = \left[-e^{-\lambda x}.e^{-\lambda a} + e^{-\lambda a}\right]$$

$$= e^{-\lambda a}\left[1 - e^{-\lambda x}\right], \text{ and}$$

$$P[X \geq a] = \int_a^{\infty} f(x)\,dx = \int_a^{\infty} \lambda e^{-\lambda x}dx = \lambda\left[\frac{e^{-\lambda x}}{-\lambda}\right]_a^{\infty} = -\left[0 - e^{-a\lambda}\right] = e^{-\lambda a}$$

$\therefore P\left[X \le x + a \mid X \ge a\right] = \dfrac{e^{-\lambda a}\left[1 - e^{-\lambda x}\right]}{e^{-\lambda a}} = 1 - e^{-\lambda x}$

Also, $P[\, X \le x \,] = \int_0^x \lambda e^{-\lambda x} dx$

$= 1 - e^{-\lambda x}$ [On simplification]

Thus,

$P[X \le x + a \mid X \ge a] = P[X \le x]$.

Hence proved

Example 4: Show that for the exponential distribution:
$f(x) = Ae^{-x}$, $0 \le x < \infty$, mean and variance are equal.

Solution: As $f(x)$ is probability function,

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

$\Rightarrow \int_0^\infty Ae^{-x} dx = 1 \Rightarrow A\left[\dfrac{e^{-x}}{(-1)}\right]_0^\infty = 1$

$\Rightarrow -A\,[0 - 1] = 1 \Rightarrow A = 1$

$\therefore f(x) = e^{-x}$

Now, comparing it with the exponential distribution

$f(x) = \lambda e^{-\lambda x}$, we have

$\lambda = 1$

Hence, mean $= \dfrac{1}{\lambda} = \dfrac{1}{1} = 1$,

and variance $= \dfrac{1}{\lambda^2} = \dfrac{1}{1} = 1$.

So, the mean and variance are equal for the given exponential distribution.

Example 5: Telephone calls arrive at a switchboard following an exponential distribution with parameter $\lambda = 12$ per hour. If we are at the switchboard, what is the probability that the waiting time for a call is

i) at least 15 minutes

ii) not more than 10 minutes.

Solution: Let X be the waiting time (in hours) for a call.

$\therefore f(x) = \lambda e^{-\lambda x}$, $x \ge 0$

$\Rightarrow F(x) = P\left[X \le x\right] = 1 - e^{-\lambda x}$ [c.d.f. of exponential distribution]

$= 1 - e^{-12x}$ … (1) [$\because \lambda = 12$]

Now,

i) P[waiting time is at least 15 minutes] = P[waiting time is at least $\frac{1}{4}$ hours]

$$= P\left[X \geq \frac{1}{4}\right] = 1 - P\left[X < \frac{1}{4}\right]$$

$$= 1 - \left[1 - e^{-12\times\frac{1}{4}}\right] \quad \text{[Using (1) above]}$$

$$= e^{-3}$$

$$= 0.0498 \quad \left[\begin{array}{l}\text{See table given at the}\\ \text{end of Unit 10}\end{array}\right]$$

ii) P[waiting time not more than 10 minutes]

= P[waiting time not more than $\frac{1}{6}$ hrs]

$$= P\left[X \leq \frac{1}{6}\right] = 1 - e^{-12\times\frac{1}{6}}$$

$$= 1 - e^{-2} = 1 - (0.1353) = 0.8647$$

Now, we are sure that you can try the following exercises.

E3) What are the mean and variance of the exponential distribution given by:

$$f(x) = 3e^{-3x}, x \geq 0$$

E4) Obtain the value of k > 0 for which the function given by

$$f(x) = 2e^{-kx}, x \geq 0$$

follows an exponential distribution.

E5) Suppose that accidents occur in a factory at a rate of $\lambda = \frac{1}{20}$ per working day. Suppose in the factory six days (from Monday to Saturday) are working. Suppose we begin observing the occurrence of accidents at the starting of work on Monday. Let X be the number of days until the first accident occurs. Find the probability that

i) first week is accident free

ii) first accident occurs any time from starting of working day on Tuesday in second week till end of working day on Wednesday in the same week.

We now conclude this unit by giving a summary of what we have covered in it.

15.4 SUMMARY

Following main points have been covered in this unit.

1) A random variable X is said to follow a **continuous uniform (rectangular)** distribution over an interval (a, b) if its probability density function is given by

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

2) For **continuous uniform distribution**, **Mean** $= \dfrac{a+b}{2}$ and

variance $= \dfrac{(b-a)^2}{12}$.

3) A random variable X is said to follow **exponential distribution** with parameter $\lambda > 0$, if it takes any non-negative real value and its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{, elsewhere} \end{cases}$$

4) For **exponential distribution**, **Mean** = $\dfrac{1}{\lambda}$ and **Variance** = $\dfrac{1}{\lambda^2}$.

5) Mean > or = or < Variance according to whether λ > or = or < 1.

6) **Exponential distribution** is the **only continuous distribution** which has the **memoryless property** given by:

$P[X \leq x + a \mid X \geq a] = P[X \leq x]$.

15.5 SOLUTIONS/ANSWERS

E1) As $X \sim U[-a, a]$,

$\therefore$ probability density function of X is

$$f(x) = \frac{1}{a-(-a)} = \frac{1}{a+a} = \frac{1}{2a}, \quad -a < x < a.$$

i) Given that $P[X > 4] = \dfrac{1}{3}$

$$\Rightarrow \int_4^a \frac{1}{2a} dx = \frac{1}{3}$$

$$\Rightarrow \frac{1}{2a}[x]_4^a = \frac{1}{3}$$

$$\Rightarrow \frac{a-4}{2a} = \frac{1}{3}$$

$$\Rightarrow 3a - 12 = 2a$$

$$\Rightarrow a = 12.$$

ii) $P[X < 1] = \frac{3}{4}$

$$\Rightarrow \int_{-a}^{1} \frac{1}{2a} dx = \frac{3}{4}$$

$$\Rightarrow \frac{1}{2a}[x]_{-a}^{1} = \frac{3}{4}$$

$$\Rightarrow \frac{1}{2a}[1+a] = \frac{3}{4}$$

$$\Rightarrow 1 + a = \frac{3}{2}a$$

$$\Rightarrow 2 + 2a = 3a$$

$$\Rightarrow a = 2$$

iii) $P[|X| < 2] = P[|X| > 2]$

$$\Rightarrow P[-2 < X < 2] = P[X < -2 \text{ or } X > 2] \quad \left[\begin{array}{l} \because |X| < 2 \Rightarrow \pm X < 2 \\ \Rightarrow X < 2 \text{ or } -X < 2 \\ \Rightarrow -2 < X < 2 \text{ and} \\ |X| > 2 \Rightarrow \pm X > 2 \\ \Rightarrow X > 2 \text{ or } -X > 2 \\ \Rightarrow X > 2 \text{ or } X < -2 \end{array}\right]$$

$$\Rightarrow P[-2 < X < 2] = P[X < -2] + P[X > 2] \quad \left[\begin{array}{l} \text{By Addition law of} \\ \text{probability for mutually} \\ \text{exclusive events} \end{array}\right]$$

$$\Rightarrow \int_{-2}^{2} \frac{1}{2a} dx = \int_{-a}^{-2} \frac{1}{2a} dx + \int_{2}^{a} \frac{1}{2a} dx$$

$$\Rightarrow \frac{1}{2a}[4] = \frac{1}{2a}[-2+a] + \frac{1}{2a}[a-2]$$

$$\Rightarrow 4 = (-2+a) + (a-2)$$

$$\Rightarrow 4 = -4 + 2a$$

$$\Rightarrow 2a = 8$$

$$\Rightarrow a = 4$$

E2) As $X \sim U[-2, 2]$,

$$\therefore f(x) = \frac{1}{4}, -2 < x < 2.$$

Now $\quad P[X > k] = \frac{1}{2}$

$\Rightarrow \int_k^2 \frac{1}{4} dx = \frac{1}{2}$

$\Rightarrow \frac{2-k}{4} = \frac{1}{2}$

$\Rightarrow 2 - k = 2$

$\Rightarrow k = 0.$

E3) Comparing it with the exponential distribution given by

$f(x) = \lambda e^{-\lambda x}, x \geq 0$

We have $\lambda = 3$

$\therefore$ Mean $= \frac{1}{\lambda} = \frac{1}{3}$ and Variance $= \frac{1}{\lambda^2} = \frac{1}{9}$

E4) As the given function is exponential distribution i.e. a p.d.f.,

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

$\Rightarrow k = 2$ [On simplification]

Alternatively, you may compare the given function with exponential distribution

$f(x) = \lambda e^{-\lambda x},$

we have

$\lambda = 2$ and $\lambda = k$

$\therefore k = 2$

E5) Here $P[X \leq x] = F(x) = 1 - e^{-\lambda x} = 1 - e^{-\frac{1}{20}x}$

i) P[First week is accident free] = P[Accident occurs after six days]

$= P[X > 6] = 1 - P[X \leq 5]$

$= 1 - \left[1 - e^{-5/20}\right] = e^{-\frac{1}{4}} = e^{-0.25} = 0.7788.$

ii) P[First accident occurs on second week from starting of working day on Tuesday till end of working day on Wednesday]

=P[First accident occurs after 7 working days and before the end of 9 working days]

$= P[7 < X \leq 9]$

$= P[X \leq 9] - P[X \leq 7]$

$$= \left(1 - e^{-\frac{9}{20}}\right) - \left(1 - e^{-\frac{7}{20}}\right)$$

$$= -e^{-\frac{9}{20}} + e^{-\frac{7}{20}}$$

$$= e^{-\frac{7}{20}} - e^{-\frac{9}{20}}$$

$$= e^{-0.35} - e^{-0.45}$$

$= 0.7047 – 0.6376$ [See the table give at the end of Unit 10]

$= 0.0671.$

UNIT 16 GAMMA AND BETA DISTRIBUTIONS

Structure

16.1 INTRODUCTION

In Unit 15, you have studied continuous uniform and exponential distributions. Here, we will discuss gamma and beta distributions. Gamma distribution reduces to exponential distribution and beta distribution reduces to uniform distribution for special cases. Gamma distribution is a generalization of exponential distribution in the same sense as the negative binomial distribution is a generalization of geometric distribution. In a sense, the geometric distribution and negative binomial distribution are the discrete analogs of the exponential and gamma distributions, respectively. The present unit discusses the gamma and beta distributions which are defined with the help of special functions known as gamma and beta functions, respectively. So, before defining these distributions, we first define gamma and beta functions in Sec. 16.2 of this unit. Then gamma distribution and beta distribution of first kind followed by beta distribution of second kind are discussed in Secs. 16.3 to 16.5.

Objectives

After studing this unit, you would be able to:

- define beta and gamma functions;
- define gamma and beta distributions;
- discuss various properties of these distributions;
- identify the situations where these distributions can be employed; and
- solve various practical problems related to these distributions.

16.2 BETA AND GAMMA FUNCTIONS

In this section, some special functions i.e. beta and gamma functions are defined with their properties and the relation between them. These will be helpful in defining beta and gamma distributions to be defined in the subsequent sections.

Beta Function

Definition: If m > 0, n > 0, the integral $\int_0^1 x^{m-1}(1-x)^{n-1}dx$ is called a beta function and is denoted by β(m, n) e.g.

i) $\int_0^1 \sqrt{x}(1-x)^2\,dx = \int_0^1 x^{\frac{3}{2}-1}(1-x)^{3-1}dx = \beta\left(\frac{3}{2},3\right)$

or $\int_0^1 \sqrt{x}(1-x)^2\,dx = \beta\left(\frac{1}{2}+1, 2+1\right) = \beta\left(\frac{3}{2},3\right)$

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

Properties of Beta Function

1. Beta function is symmetric function i.e. β(m, n) = β(n, m)
2. There are some other forms also of Beta function. One of these forms, which will be helpful in defining beta distribution of second kind, is

$$\beta(m,n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}dx$$

3. (i) $\frac{\beta(p,q+1)}{q} = \frac{\beta(p+1,q)}{p}$

   (ii) β(p, q) = β(p + q, q) × β(p, q + 1)

On the basis of the above discussion, you can try the following exercise.

E1) Express the following as a beta function:

i) $\int_0^1 x^{-\frac{1}{3}}(1-x)^{\frac{1}{2}}dx$

ii) $\int_0^1 x^{-2}(1-x)^5 dx$

iii) $\int_0^\infty \frac{x^2}{(1+x)^5}dx$

iv) $\int_0^\infty \frac{x^{-\frac{1}{2}}}{(1+x)^2}dx$

Gamma Function

Though we have defined Gamma function in Unit 13, yet we are again defining it with more properties, examples and exercises to make you clearly understand this special function.

Definition: If n > 0, the integral $\int_0^\infty x^{n-1}e^{-x}dx$ is called a gamma function and is denoted by $\Gamma(n)$

e.g.

(i) $\int_0^\infty x^2e^{-x}dx = \Gamma(2+1) = \Gamma(3)$

(ii) $\int_0^\infty \sqrt{x}e^{-x}dx = \Gamma\left(\frac{1}{2}+1\right) = \Gamma\left(\frac{3}{2}\right)$

Some Important Results on Gamma Function

1. If n > 1, $\Gamma(n) = (n-1)\Gamma(n-1)$
2. If n is a positive integer, $\Gamma n = (n-1)!$
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Relationship between Beta and Gamma Functions

If m > 0, n > 0, then $\beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$

You can now try the following exercise.

E2) Evaluate:

(i) $\int_0^\infty e^{-x}x^{\frac{5}{2}}dx$

(ii) $\int_0^\infty (1-x)^{10}dx$

(iii) $\int_0^\infty x^{-\frac{1}{2}}e^{-x}dx$

16.3 GAMMA DISTRIBUTION

Gamma distribution is a generalisation of exponential distribution. Both the distributions are good models for waiting times. For exponential distribution, the length of time interval between successive happenings is considered i.e. the time is considered till one happening occurs whereas for gamma distribution, the length of time between 0 and the instant when r^{th} happening occurs is considered. So, if r = 1, then the situation becomes the exponential situation. Let us now define gamma distribution:

Definition: A random variable X is said to follow gamma distribution with parameters r > 0 and λ > 0 if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{\lambda^r e^{-\lambda x} x^{r-1}}{\Gamma(r)}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

Remark 1:

(i) It can be verified that

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

Verification:

$$\int_0^\infty (x)dx = \int_0^\infty \frac{\lambda^r e^{-\lambda x} x^{r-1}}{\Gamma(r)} dx$$

$$= \int_0^\infty \frac{\lambda e^{-\lambda x} (\lambda x)^{r-1}}{\Gamma(r)} dx$$

Putting $\lambda x = y \Rightarrow \lambda\, dx = dy$

Also, when $x = 0, y = 0$ and when $x \to \infty, y \to \infty$

$$= \frac{1}{\Gamma(r)} \int_0^\infty e^{-y} y^{r-1} dx$$

$$= \frac{1}{\Gamma(r)} \Gamma(r) \qquad \text{[Using gamma function defined in Sec. 16.2]}$$

= 1

(ii) If X is a gamma variate with two parameters r > 0 and λ > 0, it is expressed as $X \sim \gamma(\lambda, r)$.

(iii) If we put r = 1, we have

$$f(x) = \frac{\lambda e^{-\lambda x} x^0}{\Gamma(1)}, x > 0$$

$$= \lambda e^{-\lambda x}, x > 0$$

which is probability density function of exponential distribution.

Hence, exponential distribution is a particular case of gamma distribution.

(iv) If we put λ = 1, we have

$$f(x) = \frac{e^{-x}.x^{r-1}}{\Gamma r}, x > 0, r > 0$$

It is known as gamma distribution with single parameter r. This form of the gamma distribution is also widely used. If X follows gamma distribution with single parameter r > 0, it is expressed as $X \sim \gamma(r)$.

Mean and Variance of Gamma Distribution

If X has a gamma distribution with parameters r > 0 and $\lambda > 0$, then its

Mean = $\frac{r}{\lambda}$, Variance = $\frac{r}{\lambda^2}$.

If X has a gamma distribution with single parameter r > 0, then its

Mean = Variance = r.

Additive Property of Gamma Distribution

1. If $X_1, X_2, \ldots, X_k$ are independent gamma variates with parameters $(\lambda, r_1), (\lambda, r_2), \ldots, (\lambda, r_k)$ respectively, then $X_1 + X_2 + \ldots + X_k$ is also a gamma variate with parameter $(\lambda, r_1 + r_2 + \ldots + r_k)$.
2. If $X_1, X_2, \ldots, X_k$ are independent gamma variates with single parameters $r_1, r_2, \ldots, r_k$ respectively, then $X_1 + X_2 + \ldots, + X_k$ is also a gamma variate with parameter $r_1 + r_2 + \ldots + r_k$.

Example 1: Suppose that on an average 1 customer per minute arrive at a shop. What is the probability that the shopkeeper will wait more than 5 minutes before

(i) both of the first two customers arrive, and

(ii) the first customer arrive?

Assume that waiting times follows gamma distribution.

Solution:

i) Let X denotes the waiting time in minutes until the second customer arrives, then X has gamma distribution with r = 2 (as the waiting time is to be considered up to 2nd customer)

$\lambda = 1$ customer per minute.

$$\therefore\ P[X > 5] = \int_5^\infty f(x)dx = \int_5^\infty \frac{\lambda^r e^{-\lambda x} x^{r-1}}{\Gamma(r)} dx$$

$$= \int_5^\infty \frac{(1)^2 e^{-x} x^{2-1}}{\Gamma(2)} dx = \int_5^\infty \frac{e^{-x} x^1}{\underline{|1}} dx = \int_5^\infty x^1 e^{-x} dx$$

$$= \left[\left\{ x \frac{e^{-x}}{-1} \right\}_5^\infty - \int_5^\infty (1) \frac{e^{-x}}{-1} dx \right] \qquad \text{[Integrating by parts]}$$

$$= \left(0 + 5e^{-5}\right) + \int_5^\infty e^{-x} dx = 5e^{-5} + \left[\frac{e^{-x}}{-1} \right]_5^\infty$$

$$= 5e^{-5} - \left(0 - e^{-5}\right)$$

$$= 6e^{-5}$$

$= 6 \times 0.0070$ [See the table given at the end of Unit 10]

$= 0.042$

ii) In this case r = 1, λ = 1 and hence

$$P[X > 5] = \int_5^\infty \frac{\lambda^r e^{-\lambda x}.x^{r-1}}{\overline{|(r)}} dx$$

$$= \int_5^\infty \frac{(1)^1 e^{-x} x^0}{\overline{|(1)}} dx = \int_5^\infty e^{-x} dx = \left[\frac{e^{-x}}{-1}\right]_5^\infty = 0 + e^{-5} = 0.0070$$

Alternatively,

As r = 1, so it is a case of exponential distribution for which

$f(x) = \lambda e^{-\lambda x}, x > 0$

$$\therefore P[X > 5] = \int_5^\infty \lambda e^{-\lambda x} dx = \int_5^\infty (1) e^{-x} dx = \left[\frac{e^{-x}}{-1}\right]_5^\infty = 0 + e^{-5} = 0.0070$$

Here is an exercise for you.

E3) Telephone calls arrive at a switchboard at an average rate of 2 per minute. Let X denotes the waiting time in minutes until the 4^{th} call arrives and follows gamma distribution. Write the probability density function of X. Also find its mean and variance.

Let us now discuss the beta distributions in the next two sections:

16.4 BETA DISTRIBUTION OF FIRST KIND

You have studied in Sec. 16.3 that beta function is related to gamma function in the following manner:

$$\beta(m,n) = \frac{\overline{|(m)}\,\overline{|(n)}}{\overline{|(m+n)}}$$

Now, we are in a position to define beta distribution which is defined with the help of beta function. There are two kinds of beta distribution – beta distribution of first kind and beta distribution of second kind. Beta distribution of second kind is defined in next section of the unit whereas beta distribution of first kind is defined as follows:

Definition: A random variable X is said to follow beta distribution of first kind with parameters m > 0 and n > 0, if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{\beta(m,n)} x^{m-1}(1-x)^{n-1}, 0 < x < 1 \\ 0, \qquad \text{otherwise} \end{cases}$$

The random variable X is known as beta variate of first kind and can be expressed as $X \sim \beta_1(m, n)$

Remark 5: If m = 1 and n = 1, then the beta distribution reduces to

$$f(x)=\frac{1}{\beta(1,1)}x^{1-1}(1-x)^{1-1},\ 0<x<1$$

$$=\frac{x^0(1-x)^0}{\beta(1,1)},\ 0<x<1$$

$$=\frac{1}{\beta(1,1)},0<x<1$$

But $\beta(1,1)=\frac{\Gamma 1\,\Gamma 1}{\Gamma 2}=\frac{\lfloor 0\lfloor 0}{\lfloor 1}$

Therefore, $f(x)=\frac{(1)(1)}{(1)}=1$

$\therefore f(x)=1,\ 0<x<1$

$$=\frac{1}{1-0},0<x<1$$

which is uniform distribution on (0, 1).

[$\because$ p.d.f. of uniform distribution on (a, b) is $f(x)=\frac{1}{b-a}$, $a<x<b$]

So, continuous uniform distribution is a particular case of beta distribution.

Mean and variance of Beta Distribution of First Kind

Mean and Variance of this distribution are given as

Mean $=\frac{m}{m+n}$

Variance = $\frac{mn}{(m+n)^2(m+n+1)}$

Example 4: Determine the constant C such that the function $f(x)=Cx^3(1-x)^6, 0<x<1$ is a beta distribution of first kind. Also, find its mean and variance.

Solution: As $f(x)$ is a beta distribution of first kind.

$$\therefore \int_0^1 f(x)=1$$

$$\Rightarrow \int_0^1 Cx^3(1-x)^6\,dx=1$$

$$\Rightarrow C\int_0^1 x^3(1-x)^6\,dx=1$$

$\Rightarrow\ C\beta(3+1, 6+1) = 1$ [By definition of Beta distribution of first kind]

$$\Rightarrow\ C = \frac{1}{\beta(4,7)}$$

$$= \frac{\Gamma(4+7)}{\Gamma(4)\Gamma(7)} \qquad \left[\because \beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}\right]$$

$$= \frac{\Gamma(11)}{\Gamma(4)\Gamma(7)} = \frac{\lfloor 10}{\lfloor 3 \lfloor 6}$$

$$= \frac{10\times 9\times 8\times 7\times \lfloor 6}{3\times 2\times \lfloor 6} = 840$$

Thus, $f(x) = 840x^3(1-x)^6$

$$= 840x^{4-1}(1-x)^{7-1}$$

$$= \frac{x^{4-1}(1-x)^{7-1}}{\beta(4,7)}$$

$\left[\because \frac{1}{\beta(4,7)} = 840\right.$ just obtained above in this example]

$\therefore$ m = 4, n = 7

$$\Rightarrow\ \text{Mean} = \frac{m}{m+n} = \frac{4}{4+7} = \frac{4}{11},$$

and Variance $= \dfrac{mn}{(m+n)^2(m+n+1)}$

$$= \frac{4\times 7}{(4+7)^2(4+7+1)}$$

$$= \frac{28}{(121)(12)} = \frac{7}{(121)(3)} = \frac{7}{363}$$

Now, you can try the following exercises.

E4) Using beta function, prove that

$$\int_0^1 60x^2(1-x)^3\,dx = 1$$

E5) Determine the constant k such that the function

$f(x) = kx^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}, 0 < x < 1,$ is a beta distribution of first kind. Also find its mean and variance.

16.5 BETA DISTRIBUTION OF SECOND KIND

Let us now define beta distribution of second kind.

Definition: A random variable X is said to follow beta distribution of second kind with parameters m > 0, n > 0 if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{x^{m-1}}{\beta(m,n)(1+x)^{m+n}}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

Remark 6: It can be verified that $\int_0^\infty \frac{x^{m-1}}{\beta(m,n)(1+x)^{m+n}} dx = 1$

Verification:

$$\int_0^\infty \frac{x^{m-1}}{\beta(m,n)(1+x)^{m+n}} dx = \frac{1}{\beta(m,n)} \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}} dx$$

$$= \frac{1}{\beta(m,n)} \beta(m,n)$$

$\left[\because \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}} dx \right.$ is another form of beta function. (see Sec. 16.2 of this Unit) $\Big]$

$$= 1$$

Remark 7: If X is a beta variate of second kind with parameters m > 0, n > 0, then it is expressed as $X \sim \beta_2(m, n)$

Mean and Variance of beta Distribution of second kind

Mean = $\dfrac{m}{n-1}, n > 1;$

Variance = $\dfrac{m(m+n-1)}{(n-1)^2(n-2)}, n > 2$

Example 5: Determine the constant k such that the function

$$f(x) = \frac{kx^3}{(1+x)^7}, 0 < x < \infty,$$

is the p.d.f of beta distribution of second kind. Also find its mean and variance.

Solution: As $f(x)$ is a beta distribution of second kind,

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

$$\Rightarrow \int_0^\infty \frac{kx^3}{(1+x)^7}dx = 1$$

$$\Rightarrow k\int_0^\infty \frac{x^{4-1}}{(1+x)^{4+3}}dx = 1$$

$$\Rightarrow k\beta(4,3) = 1$$

$$\Rightarrow k = \frac{1}{\beta(4,3)} = \frac{\lceil 7}{\lceil 4 \lceil 3} = \frac{\lfloor 6}{\lfloor 3 \lfloor 2} = \frac{6\times 5\times 4}{2} = 60$$

Here m = 4, n = 3

$$\therefore \text{ Mean } = \frac{m}{n-1} = \frac{4}{3-1} = \frac{4}{2} = 2$$

$$\text{Variance } = \frac{m(m+n-1)}{(n-1)^2(n-2)} = \frac{4(4+3-1)}{(3-1)^2(3-2)} = \frac{4\times 6}{4\times 1} = 6$$

Now, you can try the following exercises.

E6) Using beta function, prove that

$$\int_0^\infty \frac{x^3}{(1+x)^{\frac{13}{2}}}dx = \frac{64}{15015}$$

E7) Obtain mean and variance for the beta distribution whose density is given by

$$f(x) = \frac{60x^2}{(1+x)^7}, 0 < x < \infty$$

16.6 SUMMARY

The following main points have been covered in this unit:

1) A random variable X is said to follow **gamma distribution with parameters r > 0 and λ > 0** if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{\lambda^r e^{-\lambda x} x^{r-1}}{\lceil (r)}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

2) **Gamma distribution** of random variable X **with single parameter** r > 0 is defined as $f(x) = \frac{e^{-x}x^{r-1}}{\lceil r}, x > 0, r > 0$

3) For **gamma distribution** with two parameters λ and r, **Mean** = $\frac{r}{\lambda}$ and **Variance** = $\frac{r}{\lambda^2}$.

4) A random variable X is said to follow **beta distribution of first kind** with parameters m > 0 and n > 0, if its probability density function is given by

$$f(x)=\begin{cases}\dfrac{1}{\beta(m,n)}x^{m-1}(1-x)^{n-1}, 0<x<1\\ 0, \qquad \text{otherwise}\end{cases}$$

Its **mean and variance** are $\dfrac{m}{m+n}$ and $\dfrac{mn}{(m+n)^2(m+n+1)}$, respectively.

5) A random variable X is said to follow **beta distribution of second kind** with parameters m > 0, n > 0 if its probability density function is given by:

$$f(x)=\begin{cases}\dfrac{x^{m-1}}{\beta(m,n)(1+x)^{m+n}}, 0<x<\infty\\ 0, \qquad \text{elsewhere}\end{cases}$$

Its **Mean and Variance** are $\dfrac{m}{n-1}, n>1;$ and $\dfrac{m(m+n-1)}{(n-1)^2(n-2)}, n>2$ respectively.

6) Exponential distribution is a particular case of gamma distribution and continuous uniform distribution is a particular case of beta distribution.

16.7 SOLUTIONS/ANSWERS

E1) (i) $\int_0^1 x^{-\frac{1}{3}}(1-x)^{\frac{1}{2}}dx = B\left(-\frac{1}{3}+1, \frac{1}{2}+1\right) = B\left(\frac{2}{3}, \frac{3}{2}\right)$

(ii) $\int_0^1 x^{-2}(1-x)^5 dx = \int_0^1 x^{-1-1}(1-x)^{6-1}dx$

is not a beta function, since m = -1 < 0, but m and n both should be positive.

(iii) $\int_0^\infty \frac{x^2}{(1+x)^5}dx = \int_0^\infty \frac{x^{3-1}}{(1+x)^{3+2}}dx = \beta(3, 2)$

[$\because$ m = 3, n = 2(see Property 2 of Beta function Sec. 16.2)]

(iv) $\int_0^\infty \frac{x^{-\frac{1}{2}}}{(1+x)^2}dx = \int_0^\infty \frac{x^{\frac{1}{2}-1}}{(1+x)^{\frac{1}{2}+\frac{3}{2}}}dx = \beta\left(\frac{1}{2}, \frac{3}{2}\right)$

E2) $\int_0^\infty e^{-x}.x^{5/2}dx = \Gamma\left(\frac{5}{2}+1\right)$

$$= \Gamma\left(\frac{7}{2}\right)$$

$$= \left(\frac{5}{2}\right)\Gamma\left(\frac{5}{2}\right) = \left(\frac{5}{2}\right)\left(\frac{3}{2}\right)\Gamma\left(\frac{3}{2}\right) = \left(\frac{5}{2}\right)\left(\frac{3}{2}\right)\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)$$

$\left[\begin{array}{l}\text{Result 1 on gamma}\\ \text{function (See Sec. 16.2)}\end{array}\right]$

$$= \left(\frac{5}{2}\right)\left(\frac{3}{2}\right)\left(\frac{1}{2}\right)\sqrt{\pi} \qquad [\text{Result 3 on gamma function}]$$

$$= \left(\frac{15}{8}\right)\sqrt{\pi}$$

(ii) $\int_0^\infty x^1 (1-x)^{10}\, dx = \beta(1 + 1, 10 + 1)$

$$= \beta(2,11)$$

$$= \frac{\Gamma 2\,\Gamma 11}{\Gamma(13)} \qquad \left[\begin{array}{l}\because \text{see relation between}\\ \text{beta and gamma function}\end{array}\right]$$

$$= \frac{(1)!(10)!}{(12)!} \qquad [\text{Result 2 on gamma function}]$$

$$= \frac{(10)!}{(12)(11)(10)!} = \frac{1}{12\times 11} = \frac{1}{132}$$

(iii) $\int_0^\infty x^{-\frac{1}{2}} e^{-x} dx = \Gamma\left(-\frac{1}{2}+1\right) = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

E3) Here $\lambda = 2$, $r = 4$.

$$\therefore f(x) = \frac{\lambda^r e^{-\lambda x} . x^{r-1}}{\Gamma(r)}, x > 0$$

$$= \frac{2^4 . e^{-2x} . x^3}{\Gamma 4}, x > 0$$

$$= \frac{16 e^{-2x} . x^3}{\lfloor 3}, x > 0$$

$$= \frac{8}{3} x^3 e^{-2x}, x > 0$$

Mean = $\frac{r}{\lambda} = \frac{4}{2} = 2$,

Variance = $\frac{r}{\lambda^2} = \frac{4}{2^2} = 1$

E4) $\int_0^1 60x^2 (1-x)^3 dx = 60\int_0^1 x^{3-1} (1-x)^{4-1} dx = 60\,\beta(3, 4)$

$$= 60\frac{\Gamma(3)\Gamma(4)}{\Gamma(7)} = 60\times\frac{\lfloor 2\,\lfloor 3}{\lfloor 6} = \frac{60\times2\times3\times2}{6\times5\times4\times3\times2} = 1$$

E5) $\int_0^1 kx^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}\,dx = 1$

$$\Rightarrow k\beta\left(-\frac{1}{2}+1,\frac{1}{2}+1\right) = 1$$

$$\Rightarrow k = \frac{1}{\beta\left(\frac{1}{2},\frac{3}{2}\right)} = \frac{\Gamma(2)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)} = \frac{\lfloor 1}{\sqrt{\pi}\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} = \frac{2}{\sqrt{\pi}\sqrt{\pi}} = \frac{2}{\pi}$$

Now, as the given p.d.f. of beta distribution of first kind is

$$f(x) = \frac{2}{\pi}x^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}, 0 < x < 1$$

$$= \frac{x^{\frac{1}{2}-1}(1-x)^{\frac{3}{2}-1}}{\beta\left(\frac{1}{2},\frac{3}{2}\right)}, \ 0 < x < 1$$

$$\therefore m = \frac{1}{2}, \ n = \frac{3}{2}$$

and hence mean = $\frac{m}{m+n} = \frac{\frac{1}{2}}{\frac{1}{2}+\frac{3}{2}} = \frac{1}{4}$

Variance = $\frac{mn}{(m+n)^2(m+n+1)}$

$$= \frac{\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)}{\left(\frac{1}{2}+\frac{3}{2}\right)^2\left(\frac{1}{2}+\frac{3}{2}+1\right)} = \frac{\frac{3}{4}}{(2)^2(3)} = \frac{3}{4\times4\times3} = \frac{1}{16}$$

E6) $\int_0^\infty \frac{x^3}{(1+x)^{13/2}}\,dx = \int_0^\infty \frac{x^{4-1}}{(1+x)^{4+\frac{5}{2}}}dx$

$$= \beta\left(4,\frac{5}{2}\right) = \frac{\Gamma(4)\Gamma\left(\frac{5}{2}\right)}{\Gamma\left(4+\frac{5}{2}\right)} = \frac{\lfloor 3\,\Gamma\left(\frac{5}{2}\right)}{\Gamma\left(\frac{13}{2}\right)}$$

$$= \frac{6\times\Gamma\left(\frac{5}{2}\right)}{\frac{13}{2}.\frac{11}{2}.\frac{9}{2}.\frac{7}{2}.\frac{5}{2}\Gamma\left(\frac{5}{2}\right)} = \frac{6\times32}{13\times11\times9\times7\times5} = \frac{64}{15015}$$

E7) $f(x) = \dfrac{60x^2}{(1+x)^7}, 0 < x < \infty$

$$= \frac{60x^{3-1}}{(1+x)^{3+4}}, 0 < x < \infty$$

$$= \frac{x^{3-1}}{\beta(3,4)(1+x)^{3+4}},\ 0 < x < \infty \quad \left[\because \beta(3,4) = \frac{\Gamma 3\,\Gamma 4}{\Gamma 6} = \frac{\lfloor 2 \lfloor 3}{\lfloor 6} = \frac{1}{60}\right]$$

m =3, n = 4

Hence, mean $= \dfrac{m}{n-1} = \dfrac{3}{4-1} = 1$

Variance $= \dfrac{m(m+n-1)}{(n-1)^2(n-2)} = \dfrac{3(3+4-1)}{(4-1)^2(4-2)} = \dfrac{3\times 6}{9\times 2}$ = 1.