MODULE 4

Statistical Techniques – II

4.1 THEORY OF PROBABILITY

Here we define and explain certain terms which are used frequently.

(a) Trial and event. Let an experiment be repeated under essentially the same conditions and let it result in any one of the several possible outcomes. Then, the experiment is called a *trial* and the possible outcomes are known as *events* or *cases*.

For example: (*i*) Tossing of a coin is a trial and the turning up of head or tail is an event.

(*ii*) Throwing a die is a trial and getting 1 or 2 or 3 or 4 or 5 or 6 is an event.

(b) Exhaustive events. The total number of all possible outcomes in any trial is known as *exhaustive events* or *exhaustive cases*.

For example: (*i*) In tossing a coin, there are two exhaustive cases, head and tail.

(*ii*) In throwing a dice, there are 6 exhaustive cases, for any one of the six faces may turn up.

(*iii*) In throwing two dice, the exhaustive cases are $6 \times 6 = 6^2$ for any of the 6 numbers from 1 to 6 on one die can be associated with any of the 6 numbers on the other die.

In general, in throwing n dice, the exhaustive cases are 6^n.

(c) Favourable events or cases. The cases which entail the happening of an event are said to be *favourable* to the event. It is the total number of possible outcomes in which the specified event happens.

For example: (*i*) In throwing a die, the number of cases favourable to the appearance of a multiple of 3 are two *viz.* 3 and 6 while the number of cases favourable to the appearance of an even number are three, *viz.*, 2, 4 and 6.

(*ii*) In a throw of two dice, the number of cases favourable to getting a sum 6 is 5, *viz.*, (1, 5); (5, 1); (2, 4); (4, 2); (3, 3).

(d) Mutually exclusive events. Events are said to be *mutually exclusive* or *incompatible* if the happening of any one of them precludes (*i.e.*, rules out) the happening of all others, *i.e.*, if no two or more than two of them can happen simultaneously in the same trial.

For example: (*i*) In tossing a coin, the events head and tail are mutually exclusive, since if the outcome is head, the possibility of getting tail in the same trial is ruled out.

(*ii*) In throwing a die, all the six faces numbered, 1, 2, 3, 4, 5, 6 are mutually exclusive since any outcome rules out the possibility of getting any other.

(e) Equally likely events. Events are said to be *equally likely* if there is no reason to expect any one in preference to any other.

Scanned with CamScanner

For example: (*i*) When a card is drawn from a well shuffled pack, any card may appear in the draw so that the 52 different cases are equally likely.

(*ii*) In throwing a die, all the six faces are equally likely to come.

(*f*) Independent and dependent events. Two or more events are said to be *independent* if the happening or non-happening of any one does not depend (or is not affected) by the happening or non-happening of any other. Otherwise they are said to be *dependent*.

For example: If a card is drawn from a pack of well shuffled cards and replaced before drawing the second card, the result of the second draw is independent of the first draw. However, if the first card drawn is not replaced, then, the second draw is dependent on the first draw.

4.2 MATHEMATICAL (OR CLASSICAL) DEFINITION OF PROBABILITY

If a trial results in n *exhaustive, mutually exclusive* and *equally likely cases* and m of them are favourable to the happening of an event E, then the probability of happening of E is given by

$$p \text{ or } P(E) = \frac{\text{Favourable number of cases}}{\text{Exhaustive number of cases}} = \frac{m}{n}.$$

Note 1. Since the number of cases favourable to happening of E is m and the exhaustive number of cases in n, therefore, the number of cases unfavourable to happening of E are $n - m$.

Note 2. The probability that the event E will not happen is given by

$$q \text{ or } P(\overline{E}) = \frac{\text{Unfavourable number of cases}}{\text{Exhaustive number of cases}} = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - p$$

$$p + q = 1 \quad i.e., \quad P(E) + P(\overline{E}) = 1$$

Obviously, p and q are non-negative and cannot exceed unity, *i.e.*, $0 \le p \le 1, 0 \le q \le 1$.

Note 3. If P(E) = 1, E is called a *certain event i.e.*, the chance of its happening is cent per cent. If P(E) = 0, then E is an *impossible event*.

Note 4. If n cases are favourable to E and m cases are favourable to $\overline{E}$ (*i.e.*, unfavourable to E), then exhaustive number of cases = $n + m$.

$$P(E) = \frac{n}{n+m} \quad \text{and} \quad P(\overline{E}) = \frac{m}{n+m}$$

We say that "odds in *favour of E*" are $n : m$ and "odds *against* E" are $m : n$.

4.3. STATISTICAL (OR EMPIRICAL) DEFINITION OF PROBABILITY

If in n trials, an event E happens m times, then the probability of happening of E is given by

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

ILLUSTRATIVE EXAMPLES

Example 1. *A bag contains 7 white, 6 red and 5 black balls. Two balls are drawn at random. Find the probability that they will both be white.*

Sol. Total number of balls = 7 + 6 + 5 = 18.

Out of 18 balls, 2 can be drawn in $^{18}C_2$ ways.

$\therefore$ Exhaustive number of cases = $^{18}C_2 = \dfrac{18 \times 17}{2 \times 1} = 153$

Out of 7 white balls, 2 can be drawn in $^7C_2 = \dfrac{7 \times 6}{2 \times 1} = 21$ ways.

$\therefore$ Favourable number of cases = 21

$$\text{Probability} = \frac{21}{153} = \frac{7}{51}.$$

Example 2. *Four cards are drawn from a pack of cards. Find the probability that (i) all are diamonds, (ii) there is one card of each suit, and (iii) there are two spades and two hearts.*

Sol. 4 cards can be drawn from a pack of 52 cards in $^{52}C_4$ ways.

$\therefore$ Exhaustive number of cases $= {}^{52}C_4 = \dfrac{52 \times 51 \times 50 \times 49}{4 \times 3 \times 2 \times 1} = 270725.$

(*i*) There are 13 diamonds in the pack and 4 can be drawn out of them in $^{13}C_4$ ways.

$\therefore$ Favourable number of cases $= {}^{13}C_4 = \dfrac{13 \times 12 \times 11 \times 10}{4 \times 3 \times 2 \times 1} = 715.$

$$\text{Required probability} = \frac{715}{270725} = \frac{143}{54145} = \frac{11}{4165}.$$

(*ii*) There are 4 suits, each containing 13 cards.

$\therefore$ Favourable number of cases $= {}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1 = 13 \times 13 \times 13 \times 13.$

$$\text{Required probability} = \frac{13 \times 13 \times 13 \times 13}{270725} = \frac{2197}{20825}.$$

(*iii*) 2 spades out of 13 can be drawn in $^{13}C_2$ ways.

2 hearts out of 13 can be drawn in $^{13}C_2$ ways.

$\therefore$ Favourable number of cases $= {}^{13}C_2 \times {}^{13}C_2 = 78 \times 78$

$$\text{Required probability} = \frac{78 \times 78}{270725} = \frac{468}{20825}.$$

Example 3. *A bag contains 50 tickets numbered 1, 2, 3, ..., 50, of which five are drawn at random and arranged in ascending order of magnitude ($x_1 < x_2 < x_3 < x_4 < x_5$). What is the probability that $x_3 = 30$?*

Sol. Exhaustive number of cases $^{50}C_5$.

If $x_3 = 30$, then the two tickets with numbers x_1 and x_2 must come out of 29 tickets numbered 1 to 29 and this can be done in $^{29}C_2$ ways. The other two tickets with numbers x_4 and x_5 must come out of the 20 tickets number 31 to 50 and this can be done in $^{20}C_2$ ways.

$\therefore$ Favourable number of cases $= {}^{29}C_2 \times {}^{20}C_2$.

$$\text{Required probability} = \frac{^{29}C_2 \times {}^{20}C_2}{^{50}C_5} = \frac{551}{15134}.$$

4.4 RANDOM EXPERIMENT

Occurrences which can be repeated a number of times, essentially under the same conditions, and whose result cannot be predicted before hand are known as **random experiments**. For example, rolling of a die, tossing a coin, taking out balls from an urn.

Sample Space. Out of the several possible outcomes of a random experiment, one and only one can take place in a trial. The set of all these possible outcomes is called the **sample space** for the particular experiment and is denoted by S.

Scanned with CamScanner

For example, if a coin is tossed, the possible outcomes are H (Head) and T (Tail).

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

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

For example, if a coin is tossed and H and T denote 'Head' and 'Tail' respectively, then $S = \{H, T\}$. The two sample points are H and T.

Finite Sample Space. If the number of sample points in a sample space is finite, we call it a **finite sample space.** Here, we shall deal with finite sample spaces only.

Event. Every subset of S, the sample space, is called an event.

Since $S \subset S$, S itself is an event ; called a **certain event.**

Also, $\phi \subset S$, the null set is also an event, called an **impossible event.**

If $e \in S$, then e is called an **elementary event.** Every elementary event contains only one sample point.

4.5 AXIOMS

(*i*) With each event E (*i.e.*, a sample point) is associated a real number between 0 and 1, called the probability of that event and is denoted by P(E). Thus $0 \le P(E) \le 1$.

(*ii*) The sum of the probabilities of all simple (elementary) events constituting the sample space is 1. Thus $P(S) = 1$.

(*iii*) The probability of a compound event (*i.e.*, an event made up of two or more sample events) is the sum of the probabilities of the simple events comprising the compound event.

Thus, if there are n *equally likely* possible outcomes of a random experiment, then the sample space S contains n sample points and the probability associated with each sample point is $\frac{1}{n}$. [By Axiom (*ii*)]

Now, if an event E consists of m sample points, then the probability of E is

$$P(E) = \frac{1}{n} + \frac{1}{n} + \text{......} \; m \text{ times} = \frac{m}{n} = \frac{\text{Number of sample points in E}}{\text{Number of sample points in S}}.$$

This closely agrees with the classical definition of probability.

4.6 PROBABILITY OF THE IMPOSSIBLE EVENT IS ZERO, *i.e.*, $P(\phi) = 0$

Proof. Impossible event contains no sample point. As such, the sample space S and the impossible event ϕ are *mutually exclusive*.

$$\Rightarrow \quad S \cup \phi = S \quad \Rightarrow \quad P(S \cup \phi) = P(S)$$
$$\Rightarrow \quad P(S) + P(\phi) = P(S) \quad \Rightarrow \quad P(\phi) = 0.$$

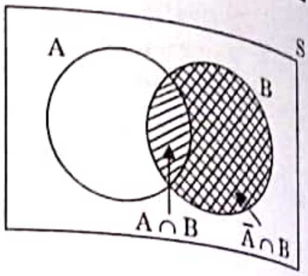
4.7. PROBABILITY OF THE COMPLEMENTARY EVENT $\bar{A}$ OF A IS GIVEN BY $P(\bar{A}) = 1 - P(A)$

Proof. $\bar{A}$ and A are disjoint events. Also $A \cup \bar{A} = S$

$$\therefore \quad P(A \cup \bar{A}) = P(S)$$
$$\Rightarrow \quad P(A) + P(\bar{A}) = 1. \text{ Hence } P(\bar{A}) = 1 - P(A).$$

Scanned with CamScanner

4.8 FOR ANY TWO EVENTS A AND B, $P(\overline{A} \cap B) = P(B) - P(A \cap B)$



Proof. $\overline{A} \cap B = \{p : p \in B \text{ and } p \notin A\}$

Now $\overline{A} \cap B$ and $A \cap B$ are disjoint sets and

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

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

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

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

Note. Similarly, it can be proved that $P(A \cap \overline{B}) = P(A) - P(A \cap B)$.

4.9 IF $B \subset A$, THEN (i) $P(A \cap \overline{B}) = P(A) - P(B)$ (ii) $P(B) \leq P(A)$

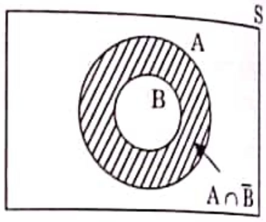
Proof. When $B \subset A$, B and $A \cap \overline{B}$ are disjoint and their union is A.

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

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

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

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



Now, if E is any event,

then $0 \leq P(E) \leq 1$, *i.e.*, $P(E) \geq 0$ [Using (1)]

$\therefore$ $P(A \cap \overline{B}) \geq 0 \Rightarrow P(A) - P(B) \geq 0$

$\Rightarrow$ $P(B) \leq P(A).$

4.10 $P(A \cap B) \leq P(A)$ AND $P(A \cap B) \leq P(B)$

Proof. By 3.9 $B \subset A \Rightarrow P(B) \leq P(A)$

Since $(A \cap B) \subset A$ and $(A \cap B) \subset B$

$\therefore$ $P(A \cap B) \leq P(A)$ and $P(A \cap B) \leq P(B)$.

4.11 ADDITION THEOREM OF PROBABILITIES (OR THEOREM OF TOTAL PROBABILITY)

Statement. *If A and B are any two events, then*

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

i.e., $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$

Proof. A and $\overline{A} \cap B$ are *disjoint sets* and their union is $A \cup B$.

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

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

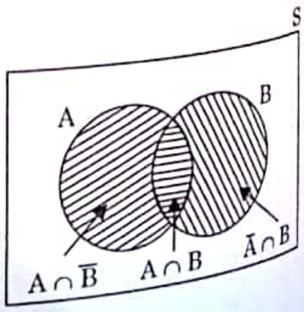
$= P(A) + [P(\overline{A} \cap B) + P(A \cap B) - P(A \cap B)]$

$= P(A) + P[(\overline{A} \cap B) \cup (A \cap B)] - P[(A \cap B)]$

[$\because$ $\overline{A} \cap B$ and $A \cap B$ are disjoint]

$= P(A) + P(B) - P(A \cap B)$ [$\because$ $(\overline{A} \cap B) \cup (A \cap B) = B$]

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



Scanned with CamScanner

Note 1. If A and B are two mutually disjoint events, then $A \cap B = \phi$, so that $P(A \cap B) = P(\phi) = 0$.

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

Note 2. $P(A \cup B)$ is also written as $P(A + B)$. Thus, for mutually disjoint events A and B,

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

$P(A \cap B)$ is also written as $P(AB)$.

4.12 IF A, B AND C ARE ANY THREE EVENTS, THEN

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

Or

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

Proof. Using the above Art. 4.11 for two events, we have

$$\begin{aligned}
P(A \cup B \cup C) &= P[(A \cup B) \cup C] \\
&= P(A \cup B) + P(C) - P[(A \cup B) \cap C] \\
&= [P(A) + P(B) - P(A \cap B)] + P(C) - P[(A \cap C) \cup (B \cap C)] \\
&\qquad\qquad\qquad\qquad [\text{By distributive law}] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - [P(A \cap C)] \\
&\qquad + P(B \cap C) - P\{(A \cap C) \cap (B \cap C)\} \qquad [\text{By } 3.11] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \\
&\qquad [\because \ (A \cap C) \cap (B \cap C) = A \cap B \cap C] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C) \\
&\qquad [\because \ A \cap C = C \cap A]
\end{aligned}$$

or $\quad P(A + B + C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC)$.

4.13 IF $A_1, A_2, ..., A_n$ ARE n MUTUALLY EXCLUSIVE EVENTS, THEN THE PROBABILITY OF THE HAPPENING OF ONE OF THEM IS

$$\mathbf{P(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)}$$

Let N be the total number of mutually exclusive, exhaustive and equally likely cases of which m_1 are favourable to A_1, m_2 are favourable to A_2 and so on.

$$\left.\begin{aligned}
&\text{Probability of occurrence of event } A_1 = P(A_1) = \frac{m_1}{N} \\
&\text{Probability of occurrence of event } A_2 = P(A_2) = \frac{m_2}{N} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\text{Probability of occurrence of event } A_n = P(A_n) = \frac{m_n}{N}
\end{aligned}\right\} \quad \ldots(1)$$

The events being mutually exclusive and equally likely, the number of cases favourable to the event

$$A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_n \text{ is } m_1 + m_2 + \ldots + m_n.$$

$\therefore$ Probability of occurrence of one of the events $A_1, A_2, \ldots\ldots, A_n$ is $P(A_1 + A_2 + \ldots + A_n)$

$$\begin{aligned}
&= \frac{m_1 + m_2 + \ldots + m_n}{N} = \frac{m_1}{N} + \frac{m_2}{N} + \ldots\ldots + \frac{m_n}{N} \\
&= P(A_1) + P(A_2) + \ldots\ldots + P(A_n) \qquad | \text{ Using (1)}
\end{aligned}$$

Scanned with CamScanner

Example 4. *In a given race, the odds in favour of four horses A, B, C, D are 1 : 3, 1 : 4, 1 : 5, 1 : 6 respectively. Assuming that a dead heat is impossible; find the chance that one of them wins the race.*

Sol. Let p_1, p_2, p_3, p_4 be the probabilities of winning of the horses A, B, C, D respectively. Since a *dead heat* (in which all the four horses cover *same distance in same time*) is not possible, the events are mutually exclusive.

Odds in favour of A are 1 : 3 $\therefore$ $p_1 = \frac{1}{1+3} = \frac{1}{4}$

Similarly, $p_2 = \frac{1}{5}, p_3 = \frac{1}{6}, p_4 = \frac{1}{7}.$

If p is the chance that one of them wins, then

$$p = p_1 + p_2 + p_3 + p_4 = \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} = \frac{319}{420}.$$

A and B are *not* mutually exclusive.

AB = the event of drawing the ace of spades

$$P(A) = \frac{13}{52}, P(B) = \frac{4}{52}, P(AB) = \frac{1}{52}$$

$$\therefore \quad P(A + B) = P(A) + P(B) - P(AB) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}.$$

4.14 CONDITIONAL PROBABILITY

The probability of the happening of an event E_1 when another event E_2 is known to have already happened is called *Conditional Probability* and is denoted by $P(E_1/E_2)$.

Mutually Independent Events. An event E_1 is said to be independent of an event E_2 if $P(E_1/E_2) = P(E_1)$.

i.e., if the probability of happening of E_1 is independent of the happening of E_2.

4.15 MULTIPLICATIVE LAW OF PROBABILITY (OR THEOREM OF COMPOUND PROBABILITY)

The probability of simultaneous occurrence of two events is equal to the probability of one of the events multiplied by the conditional probability of the other, *i.e.*, **for two events A and B,**

$$\mathbf{P(A \cap B) = P(A) \times P(B/A)}$$

where P(B/A) represents the conditional probability of occurrence of B when the event A has already happened.

Proof. Suppose a trial results in n exhaustive, mutually exclusive and equally likely outcomes, m of them being favourable to the happening of the event A.

$$\therefore \quad \text{Probability of happening of the event A} = P(A) = \frac{m}{n} \qquad ...(1)$$

Out of *m* outcomes favourable to the happening of A, let m_1 be favourable to the happening of the event B.

$$\therefore \quad \text{Conditional probability of B, given that A has happened} = P(B/A) = \frac{m_1}{m} \qquad ...(2)$$

Now out of n exhaustive, mutually exclusive and equally likely outcomes, m_1 are favourable to the happening of '*A and B*'.

Scanned with CamScanner

$\therefore$ Probability of simultaneous occurrence of A and B

$$= P(A \cap B) = \frac{m_1}{n} = \frac{m_1}{m} \times \frac{m}{n} = \frac{m}{n} \times \frac{m_1}{m}$$

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

[Using (1) and (2)]

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

Note. $P(A \cap B)$ is also written as $P(AB)$.

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

Cor. 1. Interchanging A and B

$$P(BA) = P(B) \times P(A/B)$$

or $\quad P(AB) = P(B) \times P(A/B) \quad [\because B \cap A = A \cap B]$

Cor. 2. If A and B are independent events, then $P(B/A) = P(B)$

$\therefore \quad P(AB) = P(A) \times P(B).$

Generalisation. If $A_1, A_2, ..., A_n$ are n independent events, then

$$P(A_1 A_2 ... A_n) = P(A_1) \times P(A_2) \times ...... \times P(A_n).$$

Cor. 3. If p is the chance that an event will happen in one trial then the chance that it will happen in a succession of r trials is

$$p.p ...... p \text{ . (} r \text{ times)} = p^r.$$

Cor. 4. If $p_1, p_2, ..., p_n$ are the probabilities that certain events happen, then the probabilities of their non-happening are $1 - p_1, 1 - p_2, ..., 1 - p_n$ and, therefore, the probability of all of these failing is

$$(1 - p_1)(1 - p_2)......(1 - p_n).$$

Hence the chance in which at least one of these events much happen is

$$1 - (1 - p_1)(1 - p_2) ...... (1 - p_n).$$

Example 5. *A problem in mechanics is given to three students A, B, C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively. What is the probability that the problem will be solved?*

Sol. The probabilities of A, B, C solving the problem are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$.

The probabilities of A, B, C not solving the problem are $1 - \frac{1}{2}, 1 - \frac{1}{3}, 1 - \frac{1}{4}$ *i.e.,* $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}$.

$\therefore$ The probability that the problem is not solved by any of them $= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = \frac{1}{4}$.

Hence the probability that the problem is solved by at least one of them $= 1 - \frac{1}{4} = \frac{3}{4}$.

Scanned with CamScanner

Example 6. *The odds that a book will be favourably reviewed by three independent critics are 5 to 2, 4 to 3 and 3 to 4 respectively. What is the probability that, of the three reviews, a majority will be favourable?*

Sol. Let the three critics be A, B, C. The probabilities p_1, p_2, p_3 of the book being favourably reviewed by A, B, C are $\frac{5}{7}, \frac{4}{7}, \frac{3}{7}$ respectively.

$\therefore$ The probabilities that the book is unfavourably reviewed by A, B, C are

$$1-\frac{5}{7}=\frac{2}{7},\ 1-\frac{4}{7}=\frac{3}{7},\ 1-\frac{3}{7}=\frac{4}{7}.$$

A majority will be favourable if the reviews of at least two are favourable.

(*i*) If A, B, C all review favourably, the probability is

$$\frac{5}{7}\times\frac{4}{7}\times\frac{3}{7}=\frac{60}{343} \qquad | \; p_1p_2p_3$$

(*ii*) If A, B review favourably and C reviews unfavourably, the probability is

$$\frac{5}{7}\times\frac{4}{7}\times\frac{4}{7}=\frac{80}{343} \qquad | \; p_1p_2(1-p_3)$$

(*iii*) If A, C review favourably and B reviews unfavourably, probability is

$$\frac{5}{7}\times\frac{3}{7}\times\frac{3}{7}=\frac{45}{343} \qquad | \; p_1(1-p_2)p_3$$

(*iv*) If B, C review favourably and A reviews unfavourably, the probability is

$$\frac{2}{7}\times\frac{4}{7}\times\frac{3}{7}=\frac{24}{343} \qquad | \; (1-p_1)p_2p_3$$

Hence the probability that a majority will be favourable is

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

Example 7. *A can hit a target 4 times in 5 shots; B 3 times in 4 shots; C twice in 3 shots. They fire a volley. What is the probability that at least two shots hit?*

Sol. Probability of A's hitting the target $=\frac{4}{5}$

Probability of B's hitting the target $=\frac{3}{4}$

Probability of C's hitting the target $=\frac{2}{3}$.

For at least two hits, we may have

(*i*) A, B, C all hit the target, the probability for which is $\frac{4}{5}\times\frac{3}{4}\times\frac{2}{3}=\frac{24}{60}$.

(*ii*) A, B hit the target and C misses it, the probability for which is

$$\frac{4}{5}\times\frac{3}{4}\times\left(1-\frac{2}{3}\right)=\frac{4}{5}\times\frac{3}{4}\times\frac{1}{3}=\frac{12}{60}.$$

(*iii*) A, C hit the target and B misses it, the probability for which is

$$\frac{4}{5}\times\left(1-\frac{3}{4}\right)\times\frac{2}{3}=\frac{4}{5}\times\frac{1}{4}\times\frac{2}{3}=\frac{8}{60}.$$

Scanned with CamScanner

(*iv*) B, C hit the target and A misses it, the probability for which is

$$\left(1-\frac{4}{5}\right)\times\frac{3}{4}\times\frac{2}{3}=\frac{1}{5}\times\frac{3}{4}\times\frac{2}{3}=\frac{6}{60}.$$

Since these are mutually exclusive events,

required probability $=\frac{24}{60}+\frac{12}{60}+\frac{8}{60}+\frac{6}{60}=\frac{50}{60}=\frac{5}{6}.$

Example 8. *A has 2 shares in a lottery in which there are 3 prizes and 5 blanks ; B has 3 shares in a lottery in which there are 4 prizes and 6 blanks. Show that A's chance of success is to B's as 27 : 35.*

Sol. A can draw two tickets (out of 3 + 5 = 8) in ${}^8C_3 = 28$ ways.

A will get the blanks in ${}^5C_2 = 10$ ways. $\therefore$ A can win a prize in 28 – 10 = 18 ways

Hence A's chance of success $=\frac{18}{28}=\frac{9}{14}$

B can draw 3 tickets in ${}^{10}C_3 = 120$ ways ; B will get all blanks in ${}^6C_3 = 20$ ways.

$\therefore$ B can win a prize in 120 – 20 = 100 ways.

Hence B's chance of success $=\frac{100}{120}=\frac{5}{6}.$

$\therefore$ A's chance : B's chance $=\frac{9}{14}:\frac{5}{6}=27:35.$

Example 9. *A and B throw alternately with a single die, A having the first throw. The person who first throws ace is to win. What are their respective chances of winning?*

Sol. The chance of throwing an ace with a single die $=\frac{1}{6}$

The chance of not throwing an ace with a single die $=1-\frac{1}{6}=\frac{5}{6}$.

If A is to win, he should throw an ace in the first or third or fifth, ..., throws.

If B is to win, he should throw an ace in the second or fourth or sixth, ..., throws.

The chances that an ace is thrown in the first, second, third, ..., throws are

$$\frac{1}{6},\frac{5}{6}\cdot\frac{1}{6},\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6},\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\ \ldots\ldots \quad \text{or} \quad \frac{1}{6},\frac{5}{6}\cdot\frac{1}{6},\left(\frac{5}{6}\right)^2\cdot\frac{1}{6},\left(\frac{5}{6}\right)^3\cdot\frac{1}{6},\ \ldots\ldots$$

$$\therefore \text{ A's chance} = \frac{1}{6}+\left(\frac{5}{6}\right)^2\cdot\frac{1}{6}+\left(\frac{5}{6}\right)^4\cdot\frac{1}{6}+\ldots\ldots=\frac{\frac{1}{6}}{1-\left(\frac{5}{6}\right)^2}=\frac{6}{11}$$

Sum of an infinite G.P. $=\frac{a}{1-r}$

B's chance $=1-\frac{6}{11}=\frac{5}{11}.$

Example 10. *Cards are dealt one by one from a well-shuffled pack until an ace appears. Show that the probability that exactly n cards are dealt before the first ace appears is* $\frac{4(51-n)(50-n)(49-n)}{52\cdot 51\cdot 50\cdot 49}$.

Scanned with CamScanner

Sol. Let A be the event of drawing n non-ace cards and B, the event of drawing an ace in the $(n + 1)^{th}$ draw.

Consider the event A

n cards can be drawn out of 52 cards in $^{52}C_n$ ways.

$\Rightarrow$ Exhaustive cases = $^{52}C_n$

n non-ace cards can be drawn out of 52 cards in $^{48}C_n$ ways.

$\Rightarrow$ Favourable cases = $^{48}C_n$

$$\therefore \quad P(A) = {}^{48}C_n/{}^{52}C_n = \frac{48!}{(48-n)!\,n!} \times \frac{(52-n)!\,(n)!}{52!}$$

$$= \frac{48!.(52-n)(51-n)(50-n)(49-n)(48-n)!}{(48-n)!\cdot 52\cdot 51\cdot 50\cdot 49\cdot (48)!} = \frac{(52-n)(51-n)(50-n)(49-n)}{52\cdot 51\cdot 50\cdot 49}.$$

Consider the event B

n cards have already been drawn in the first n draws.

Exhaustive cases $= {}^{52-n}C_1 = 52 - n$; Favourable cases $= {}^4C_1 = 4$

$$\therefore \qquad P(B/A) = \frac{4}{52-n}$$

Reqd. Probability = P(A) · P(B/A)

$$= \frac{(52-n)(51-n)(50-n)(49-n)}{52\cdot 51\cdot 50\cdot 49} \times \frac{4}{52-n} = \frac{4(51-n)(50-n)(49-n)}{52\cdot 51\cdot 50\cdot 49}.$$

Example 11. *An urn contains 10 white and 3 black balls, while another urn contains 3 white and 5 black balls. Two balls are drawn from the first urn and put into the second urn and then a ball is drawn from the latter. What is the probability that it is a white ball?*

Sol. The two balls drawn from the first urn may be

(*i*) both white (*ii*) both black (*iii*) one white and one black.

Let these events be denoted by A, B, C respectively.

$$P(A) = \frac{^{10}C_2}{^{13}C_2} = \frac{10\times 9}{13\times 12} = \frac{15}{26}; \qquad P(B) = \frac{^{3}C_2}{^{13}C_2} = \frac{3\times 2}{13\times 12} = \frac{1}{26}$$

$$P(C) = \frac{^{10}C_1 \times {}^3C_1}{^{13}C_2} = \frac{10\times 3}{\frac{13\times 12}{2\times 1}} = \frac{10}{26}$$

When two balls are transferred from first urn to second urn, the second urn will contain.

(*i*) 5 white and 5 black balls (*ii*) 3 white and 7 black balls

(*iii*) 4 white and 6 black balls.

Let W denote the event of drawing a white ball from the second urn in the three cases (*i*), (*ii*) and (*iii*).

Now, $P(W/A) = \frac{5}{10}$, $P(W/B) = \frac{3}{10}$, $P(W/C) = \frac{4}{10}$

$\therefore$ Reqd. probability $= P(A) . P(W/A) + P(B) . P(W/B) + P(C) . P(W/C)$

$$= \frac{15}{26}\cdot\frac{5}{10} + \frac{1}{26}\cdot\frac{3}{10} + \frac{10}{26}\cdot\frac{4}{10} = \frac{75+3+40}{260} = \frac{118}{260} = \frac{59}{130}.$$

Scanned with CamScanner

TEST YOUR KNOWLEDGE

1. In a class of 10 students, 4 are boys and the rest are girls. Find the probability that a student selected will be a girl.
2. What is the chance that a (*i*) non-leap year (*ii*) leap year should have fifty three Sundays?
3. A card is drawn from an ordinary pack and a gambler bets that it is a spade or an ace. What are the odds against his winning the bet?
4. An integer is chosen at random from the first two hundred positive integers. What is the probability that the integer chosen is divisible by 6 or 8?
5. Six cards are drawn at random from a pack of 52 cards. What is the probability that 3 will be red and 3 black?
6. From a set of raffle tickets numbered 1 to 100, three are drawn at random. What is the probability that all are odd numbered?
7. (*a*) If from a lottery of 30 tickets, marked, 1, 2, 3, ..., 30, four tickets be drawn, what is the chance that those marked 1 and 2 are among them?

   (*b*) An urn contains 5 red and 10 black balls. Eight of them are placed in another urn. What is the chance that the latter then contains 2 red and 6 black balls?
8. A party of n persons sit at a round table. Find the odds against two specified individuals sitting next to each other.
9. A five-figured number is formed by the digits 0, 1, 2, 3, 4 (without repetition). Find the probability that the number formed is divisible by 4.
10. Three newspapers A, B, C are published in a city and a survey of readers indicates the following:

    20% read A, 16% read B, 14% read C,

    8% read both A and B, 5% read both A and C,

    4% read both B and C, 2% read all the three.

    For a person chosen at random, find the probability that he reads none of the papers.
11. A problem in Statistics is given to five students. Their chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{4}$ and $\frac{1}{5}$. What is the probability that the problem will be solved?
12. A can hit a target 5 times in 6 shots, B 4 times in 5 shots and C 3 times in 4 shots. They fire a volley. What is the probability that two shots, at least hit the target?
13. Three groups of children contain respectively 3 girls and 1 boy; 2 girls and 2 boys; 1 girl and 3 boys. One child is selected at random from each group. Show that the chance that the three selected consist of 1 girl and 2 boys is $\frac{13}{32}$.
14. Four persons are chosen at random from a group containing 3 men, 2 women and 4 children. Show that the chance that exactly two of them will be children is $\frac{5}{21}$.
15. A bag contains 10 balls, two of which are red, three blue and five black. Three balls are drawn at random from the bag. What is the probability that

    (*i*) the three balls are of different colours, (*ii*) two balls are of the same colour,

    (*iii*) the balls are all of the same colour.
16. It is 8 : 5 against a person who is 40 years old living till he is 70 and 4 : 3 against a person now 50 living till he is 80. Find the probability that one at least of these persons will be alive 30 years hence.
17. Find the chance of throwing 5 or 6 at least once in four throws of a die.
18. A has 3 shares in a lottery where there are 3 prizes and 6 blanks. B has one share in another, where there is just one prize and two blanks. Show that A has a better chance of winning a prize than B in the ratio 16 : 7.

Scanned with CamScanner

19. A, B and C, in order, toss a coin. The first one to throw a head wins. If A starts, find their respective chances of winning.
20. A speaks truth in 60% cases and B in 70% cases. In what percentages of cases are they likely to contradict each other in stating the same fact?
21. A and B throw alternately with a pair of ordinary dice. A wins if he throws 6 before B throws 7 and B wins if he throws 7 before A throws 6. If A begins, find their respective chances of winning. (Huyghen's Problem)
22. (*a*) Two cards are randomly drawn from a deck of 52 cards and thrown away. What is the probability of drawing an ace in a single draw from the remaining 50 cards ?

    (*b*) A box A contains 2 white and 4 black balls. Another box B contains 5 white and 7 black balls. A ball is transferred from the box A to the box B, then a ball is drawn from the box B. Find the probability that it is white.
23. Of the cigarette smoking population, 70% are men and 30% women, 10% of these men and 20% of these women smoke 'WILLS'. What is the probability that a person seen smoking a 'WILLS' will be a man?
24. A committee consists of 9 students two of which are from 1st year, three from 2nd year and four from 3rd year. Three students are to be removed at random. What is the chance that

    (*i*) the three students belong to different classes,

    (*ii*) two belong to the same class and third to the different class, and

    (*iii*) the three belong to the same class?
25. Five men in a company of twenty are graduates. If 3 men and picked out of 20 at random, what is the probability that

    (*i*) they are all graduates? (*ii*) at least one is graduate?
26. If A, B, C are events such that

$$P(A) = 0.3, \quad P(B) = 0.4, \quad P(C) = 0.8,$$
$$P(A \cap B) = 0.08, \quad P(A \cap C) = 0.28, \quad P(A \cap B \cap C) = 0.09$$

    If $P(A \cup B \cup C) \geq 0.75$, then show that $0.23 \leq P(B \cap C) \leq 0.48$.
27. For two events A and B, let $P(A) = 0.4$, $P(B) = p$ and $P(A \cup B) = 0.6$

    (*i*) Find p so that A and B are independent events.

    (*ii*) For what value of p are A and B are mutually exclusive?
28. A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is $\frac{1}{7}$ and that of wife's selection is $\frac{1}{5}$. What is the probability that

    (*i*) both of them will be selected, (*ii*) only one of them will be selected, and

    (*iii*) none of them will be selected?
29. Two dice are tossed once. Find the probability of getting an even number on the first die or a total of 8.
30. A drawer contains 50 bolts and 150 nuts. Half of the bolts and half of the nuts are rusted. If one item is chosen at random, what is the probability that it is rusted or is a bolt?
31. A purse contains 2 silver and 4 copper coins. A second purse contains 4 silver and 3 copper coins. If a coin is pulled out at random from one of the two purses, what is the probability that it is a silver coin?
32. A man wants to marry a girl having qualities: white complexion—the probability of getting such a girl is one in twenty; handsome dowry—the probability of getting this is one in fifty; westernised manners and etiquettes—the probability here is one in hundred. Find out the probability of his getting married to such a girl when the possession of these three attributes is independent.

33. A class consists of 80 students, 25 of them are girls and 55 boys, 10 of them are rich and the remaining poor, 20 of them are fair complexioned. What is the probability of selecting a fair complexioned rich girl?

34. Of the students attending a lecture, 50% could not see what was written on the board and 40% could not hear what the lecturer was saying. Most unfortunate 30% fell into both of these categories. What is the probability that a student picked at random was able to see and hear satisfactorily?

35. The probabilities of A, B, C solving a problem are $\frac{1}{3}, \frac{2}{7}$ and $\frac{3}{8}$ respectively. If all the three try to solve the problem simultaneously, find the probability that exactly one of them will solve it.

36. A student takes his examination in four subjects $\alpha, \beta, \gamma, \delta$. He estimates his chance of passing in α as $\frac{4}{5}$, in β as $\frac{3}{4}$, in γ as $\frac{5}{6}$ and in δ as $\frac{2}{3}$. To qualify he must pass in α and at least two other subjects. What is the probability that he qualifies?

37. A candidate is called for an interview by three companies. For the first company, there are 12 candidates and for the second, there are 15 candidates and for the third, there are 10 candidates. What are the chances of his getting at least one of the companies? *(G.B.T.U. 2012)*

Answers

1. $\frac{3}{5}$
2. (i) $\frac{1}{7}$ (ii) $\frac{2}{7}$
3. 9 : 4
4. $\frac{1}{4}$
5. $\frac{13000}{39151}$
6. $\frac{4}{33}$
7. (a) $\frac{2}{145}$, (b) $\frac{140}{429}$
8. $(n - 3) : 2$
9. $\frac{5}{16}$
10. $\frac{13}{20}$
11. $\frac{17}{20}$
12. $\frac{107}{120}$
15. (i) $\frac{1}{4}$ (ii) $\frac{79}{120}$ (iii) $\frac{11}{120}$
16. $\frac{59}{91}$
17. $\frac{65}{81}$
19. $\frac{4}{7}, \frac{2}{7}, \frac{1}{7}$
20. 46%
21. $\frac{30}{61}, \frac{31}{61}$
22. (a) $\frac{1}{13}$, (b) $\frac{16}{39}$
23. $\frac{7}{13}$
24. (i) $\frac{2}{7}$ (ii) $\frac{55}{84}$ (iii) $\frac{5}{84}$
25. (i) $\frac{1}{114}$ (ii) $\frac{137}{228}$
27. (i) $\frac{1}{3}$ (ii) 0.2
28. (i) $\frac{1}{35}$ (ii) $\frac{2}{7}$ (iii) $\frac{24}{35}$
29. $\frac{5}{9}$
30. $\frac{5}{8}$
31. $\frac{19}{42}$
32. 0.000001
33. $\frac{5}{512}$
34. $\frac{2}{5}$
35. $\frac{25}{56}$
36. $\frac{61}{90}$.
37. 0.23

4.16 BAYE'S THEOREM *(G.B.T.U. 2011)*

If $E_1, E_2, \ldots\ldots, E_n$ are mutually exclusive and exhaustive events with $P(E_i) \neq 0$, $(i = 1, 2, \ldots\ldots, n)$ of a random experiment then for any arbitrary event A of the sample space of the above experiment with $P(A) > 0$, we have

$$P(E_i / A) = \frac{P(E_i)P(A/E_i)}{\sum_{i=1}^{n} P(E_i)P(A/E_i)}.$$

Scanned with CamScanner

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

The events $E_1, E_2, \ldots\ldots, E_n$ being exhaustive

$$S = E_1 \cup E_2 \cup \ldots\ldots \cup E_n$$

$$\therefore \quad A = A \cap S \qquad [\because A \subset S]$$

$$= A \cap (E_1 \cup E_2 \cup \ldots\ldots \cup E_n)$$

$$= (A \cap E_1) \cup (A \cap E_2) \cup \ldots\ldots \cup (A \cap E_n) \qquad \text{[Distributive Law]}$$

$$\Rightarrow \quad P(A) = P(A \cap E_1) + P(A \cap E_2) + \ldots\ldots + P(A \cap E_n)$$

$$= P(E_1)P(A/E_1) + P(E_2)P(A/E_2) + \ldots\ldots + P(E_n)P(A/E_n)$$

$$= \sum_{i=1}^{n} P(E_i)P(A/E_i) \qquad \ldots(1)$$

Now, $\quad P(A \cap E_i) = P(A)P(E_i/A)$

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

Note. The significance of Baye's Theorem may be understood in the following manner:

$P(E_i)$ is the probability of occurrence of E_i. The experiment is performed and we are told that the event A has occurred. With this information, the probability $P(E_i)$ is changed to $P(E_i/A)$. Baye's Theorem enables us to evaluate $P(E_i/A)$ if all the $P(E_i)$ and the conditional probabilities $P(A/E_i)$ are known.

IllUSTRATIVE EXAMPLES

Example 1. *A bag X contains 2 white and 3 red balls and a bag Y contains 4 white and 5 red balls. One ball is drawn at random from one of the bags and is found to be red. Find the probability that it was drawn from bag Y.*

Sol. Let $\quad E_1$: the ball is drawn from bag X; E_2: the ball is drawn from bag Y

and $\quad$ A: the ball is red.

We have to find $P(E_2/A)$. By Baye's Theorem,

$$P(E_2/A) = \frac{P(E_2)P(A/E_2)}{P(E_1)P(A/E_1) + P(E_2)P(A/E_2)} \qquad \ldots(1)$$

Since the two bags are equally likely to be selected, $P(E_1) = P(E_2) = \frac{1}{2}$

Also $\quad P(A/E_1) = P(\text{a red ball is drawn from bag X}) = \frac{3}{5}$

$$P(A/E_2) = P(\text{a red ball is drawn from bag Y}) = \frac{5}{9}$$

$\therefore$ From (1), we have $P(E_2/A) = \dfrac{\frac{1}{2} \times \frac{5}{9}}{\frac{1}{2} \times \frac{3}{5} + \frac{1}{2} \times \frac{5}{9}} = \dfrac{25}{52}$.

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

Scanned with CamScanner

Sol. Let E_1, E_2 and E_3 denote the events that a bolt selected at random is manufacture by the machines A, B and C respectively and let H denote the event of its being defective. Then

$$P(E_1) = 0.25,\ P(E_2) = 0.35,\ P(E_3) = 0.40$$

The probability of drawing a defective bolt manufactured by machine A is

$$P(H/E_1) = 0.05$$

Similarly,

$$P(H/E_2) = 0.04 \text{ and } P(H/E_3) = 0.02$$

By Baye's Theorem, we have

$$P(E_2/H) = \frac{P(E_2)P(H/E_2)}{P(E_1)P(H/E_1) + P(E_2)P(H/E_2) + P(E_3)P(H/E_3)}$$

$$= \frac{0.35 \times 0.04}{0.25 \times 0.05 + 0.35 \times 0.04 + 0.40 \times 0.02} = \frac{0.0140}{0.0345} = 0.41.$$

Example 3. *The contents of urns I, II and III are as follows:*

1 white, 2 black and 3 red balls,

2 white, 1 black and 1 red balls, and

4 white, 5 black and 3 red balls.

One urn is chosen at random and two balls drawn. They happen to be white and red. What is the probability that they come from urns I, II or III?

Sol. Let E_1 : urn I is chosen ; E_2 : urn II is chosen ; E_3 : urn III is chosen

and A : the two balls are white and red.

We have to find $P(E_1/A)$, $P(E_2/A)$ and $P(E_3/A)$.

Now $P(E_1) = P(E_2) = P(E_3) = \frac{1}{3}$

$$P(A/E_1) = P \text{ (a white and a red ball are drawn from urn I)} = \frac{{}^1C_1 \times {}^3C_1}{{}^6C_2} = \frac{1}{5}$$

$$P(A/E_2) = \frac{{}^2C_1 \times {}^1C_1}{{}^4C_2} = \frac{1}{3}\ ;\ P(A/E_3) = \frac{{}^4C_1 \times {}^3C_1}{{}^{12}C_2} = \frac{2}{11}$$

By Baye's Theorem, we have

$$P(E_1/A) = \frac{P(E_1)P(A/E_1)}{P(E_1)P(A/E_1) + P(E_2)P(A/E_2) + P(E_3)P(A/E_3)} = \frac{\frac{1}{3} \times \frac{1}{5}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{11}} = \frac{33}{118}$$

Similarly, $P(E_2/A) = \frac{55}{118}$, $P(E_3/A) = \frac{15}{59}$.

Scanned with CamScanner

ASSIGNMENT

1. Two urns contain 4 white, 6 blue and 4 white, 5 blue balls respectively. One of the urns is selected at random and a ball is drawn from it. If the ball drawn is white, find the probability that it is drawn from the
   (*i*) first urn (*ii*) second urn.
2. Three urns contain 6 red, 4 black; 4 red, 6 black and 5 red, 5 black balls respectively. One of the urns is selected at random and a ball is drawn from it. If the ball drawn is red, find the probability that it is drawn from the first urn.
3. A factory has two machines A and B. Past record shows that machine A produced 60% of the items of output and machine B produced 40% of the items. Further, 2% of the items produced by machine A were defective and 1% produced by machine B were defective. If a defective item is drawn at random, what is the probability that it was produced by machine A?
4. An insurance company insured 2000 scooter drivers, 4000 car drivers and 6000 truck drivers. The probability of accident is 0.01, 0.03 and 0.15 respectively. One of the insured persons meets an accident. What is the probability that he is a scooter driver?
5. A company has two plants to manufacture scooters. Plant I manufactures 70% of scooters and plant II manufactures 30%. At plant I, 80% of the scooters are rated standard quality and at plant II, 90% of the scooters are rated standard quality. A scooter is chosen at random and is found to be of standard quality. What is the chance that it has come from plant II?

Answers

1. (*i*) $\frac{9}{19}$ (*ii*) $\frac{10}{19}$
2. $\frac{2}{5}$
3. $\frac{3}{4}$
4. $\frac{1}{52}$
5. $\frac{27}{83}$.

4.17 RANDOM VARIABLE

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

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

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

4.17.1 Definition

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

1. The set $[X(s) \leq x]$ is an event for any real number x.

2. There corresponds a well defined unique probability of the events, "X assumes the value x" and "X assumes any value in the interval".

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

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

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

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

In this example, we have

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

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

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

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

4.18 TYPES OF RANDOM VARIABLES

There are two types of random variables:

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

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

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

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

4.19 PROBABILITY FUNCTION (OR PROBABILITY MASS FUNCTION)

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

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

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

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

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

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

(*i*) $p(x_i) \geq 0$, (*ii*) $\Sigma\, p(x_i) = 1$.

Example. *A random variable X takes values 1, 2, 3, ... with probability mass function* $\frac{\lambda^r}{r!}$, $r = 1, 2, 3, \ldots \infty$. *Find the value of* λ.

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

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

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

$$\Rightarrow \quad e^{\lambda} - 1 = 1$$

$$\Rightarrow \quad e^{\lambda} = 2$$

$$\lambda = \log_e 2.$$

4.20 PROBABILITY DISTRIBUTION OF A DISCRETE RANDOM VARIABLE

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

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

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

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

This probability distribution may also be expressed as follows:

X:	1	2	3	4	5	6
$p(x) = P(X = x)$:	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

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

$X = x$:	0	1	2	3
$P(X = x) = p(x)$:	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

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

4.21 MEAN AND VARIANCE OF RANDOM VARIABLES

Let

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

be a discrete probability distribution.

Scanned with CamScanner

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

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

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

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

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

ILLUSTRATIVE EXAMPLES

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

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

Number of defective bulbs = 5
Number of good bulbs = 20
Total number of bulbs = 25

P(X = 0) = P (no defective) = P (all 4 good ones)

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

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

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

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

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

∴ The probability distribution of the random variable X is

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

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

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

Probability of success $= \dfrac{2}{6} = \dfrac{1}{3}$; Probability of failure $= 1 - \dfrac{1}{3} = \dfrac{2}{3}$

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

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

Scanned with CamScanner

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

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

∴ The probability distribution of the random variable X is

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

To find the mean and variance

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

Mean $\mu = \Sigma p_i x_i = 1$

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

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

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

(i) Find k, *(ii) Evaluate* $P(X < 6)$, $P(X \ge 6)$, $P(3 < X \le 6)$

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

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

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

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

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

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

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

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

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

Scanned with CamScanner

(iii) $P(X \le 1) = k = \frac{1}{10} < \frac{1}{2}$; $\quad P(X \le 2) = k + 2k = \frac{3}{10} < \frac{1}{2}$

$P(X \le 3) = k + 2k + 2k = \frac{5}{10} = \frac{1}{2}$; $\quad P(X \le 4) = k + 2k + 2k + 3k = \frac{8}{10} > \frac{1}{2}$

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

TEST YOUR KNOWLEDGE

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

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

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

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

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

Answers

1.

X :	0	1	2	3	4
P(X) :	$\frac{625}{1296}$	$\frac{500}{1296}$	$\frac{150}{1296}$	$\frac{20}{1296}$	$\frac{1}{1296}$

2.

X :	0	1	2
P(X) :	$\frac{12}{22}$	$\frac{9}{22}$	$\frac{1}{22}$

3. $\frac{4}{9}$ 4. $\frac{400}{2873}$ 5. $\frac{9}{7}, \frac{36}{49}, \frac{6}{7}$

6. (i) $a = \frac{1}{81}$ (ii) $\frac{1}{9}, \frac{8}{9}, \frac{7}{27}$ (iii) 5 7. $2\sqrt{5}$.

Scanned with CamScanner

4.22 CONTINUOUS DISTRIBUTIONS

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

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

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

4.23 PROBABILITY DENSITY FUNCTION

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

Symbolically it is expressed as

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

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

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

The density function possesses the following two properties:

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

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

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

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

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

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

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

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

Scanned with CamScanner

Hence a probability density function can be formed as given by

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

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

ILLUSTRATIVE EXAMPLES

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

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

(*a*) $f(x) \ge 0$ for every x (*b*) $\int_{-\infty}^{+\infty} f(x)\,dx = 1.$

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

The second condition will be satisfied if

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

i.e., if $$\left[-ce^{-x}\right]_0^\infty = 1$$

i.e., if $$c = 1.$$

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

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

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

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

4.24 VARIOUS MEASURES FOR CONTINUOUS PROBABILITY DISTRIBUTIONS

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

(*i*) **Arithmetic Mean** $= \int_a^b x\,f(x)\,dx$ ...(1)

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

Scanned with CamScanner

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

(*iv*) v_r (about origin) $= \int_a^b x^r . f(x)\, dx$ ...(2)

μ_r' (about the point $x = A$) $= \int_a^b (x - A)^r . f(x)\, dx$ ...[2(*a*)]

and μ_r (about mean) $= \int_a^b (x - \text{mean})^r . f(x)\, dx$ ...[2(*b*)]

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

$$\nu_1 \text{ (about origin)} = Mean = \int_a^b x\, f(x)\, dx \quad \text{and} \quad \nu_2 = \int_a^b x^2\, f(x)\, dx$$

Hence $\mu_2 = (\nu_2 - \nu_1^2) = \int_a^b x^2 f(x)\, dx - \left(\int_a^b x\, f(x)\, dx\right)^2$ ...[2(*c*)]

From (2), on putting $r = 3$ and 4 respectively, we get the values of ν_3 and ν_4 and consequently the moments about mean can be obtained and hence β_1 and β_2 can be computed.

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

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

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

for M, we get the value of median.

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

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

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

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

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

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

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

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

Scanned with CamScanner

4.25 CUMULATIVE DISTRIBUTION FUNCTION

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

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

and
$$F(b) - F(a) = \int_a^b f(x)\, dx = P(a \le X \le b).$$

$F(x)$ is called the cumulative distribution function of x or simply the *distribution function*.

The cumulative distribution function has the following properties:

(1) $F'(x) = f(x) \ge 0$, so that $F(x)$ is a non decreasing function. This means that $dF(x) = f(x)\, dx$. This is known as *probability differential of* X.

(2) $F(-\infty) = 0$

(3) $F(\infty) = 1$.

(4) $F(x)$ is a continuous function of x on the right.

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

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

find the corresponding density function of random variable X.

Sol. The probability density function $= f(x) = \dfrac{d}{dx}\,\{F(x)\}$

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

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

Example 4. *A random variable x has the density function*

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

Determine k and the distribution function.

Sol. It will be a density function if

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

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

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

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

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

Scanned with CamScanner

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

(*i*) *Check that the above is a p.d.f.,*

(*ii*) *Obtain an expression for the c.d.f. of X.,*

(*iii*) *Compute* $P\left(X \le \frac{1}{2} \middle| \frac{1}{3} \le X \le \frac{2}{3}\right)$, *and*

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

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

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

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

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

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

or $$3k^2 - 2k^3 = 3(1 - k^2) - 2\,(1 - k^3)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$F(x) = \int_{-\infty}^{0} 0\,.\,dt + \int_0^1 \frac{t}{2}dt + \int_1^2 \frac{1}{2}dt + \int_2^x \left(-\frac{t}{2} + \frac{3}{2}\right)dt.$$

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

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

$$F(x) = \int_{-\infty}^{0} 0.\,dt + \int_0^1 \frac{t}{2}\,dt + \int_1^2 \frac{1}{2}dt + \int_2^3 \left(-\frac{t}{2} + \frac{3}{2}\right)dt + \int_3^x 0.\,dt$$

$$= \frac{1}{4} + \left(1 - \frac{1}{2}\right) + \left(-\frac{9}{4} + \frac{9}{2} + 1 - 3\right) = 1$$

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

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

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

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

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

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

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

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

Sol. Let the capacity of the tank (in '000 of litres) be 'a' such that

$$P(X \geq a) = 0.01$$

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

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

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

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

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

Hence the capacity of the tank = 0.6019 × 1,000 litres = 601.9 litres.

Example 8. *A continuous random variable X has a p.d.f. $f(x) = 3x^2$, $0 \leq x \leq 1$. Find a and b such that (i) $P(X \leq a) = P(X > a)$, and (ii) $P(X > b) = 0.05$.*

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

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

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

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

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

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

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

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

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

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

Scanned with CamScanner

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

$$\text{Mean} = \mu_1' = \int_{-\infty}^{\infty} x\, f(x)\, dx = 6\int_0^1 (x^3 - x^6)\, dx = 6\left|\frac{x^4}{4} - \frac{x^7}{7}\right|_0^1 = \frac{9}{14}$$

$$\mu_2' = \int_{-\infty}^{\infty} x^2\, f(x)\, dx = 6\int_0^1 (x^4 - x^7)\, dx = 6\left|\frac{x^5}{5} - \frac{x^8}{8}\right|_0^1 = \frac{9}{20}$$

$$\text{Variance} = \mu_2 = \mu_2' - \mu_1'^2 = \left\{\frac{9}{20} - \left(\frac{9}{14}\right)^2\right\} = \frac{9}{245}.$$

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

$$\text{Mean} = \mu_1' = \int_{-\infty}^{\infty} x f(x)\, dx = k\int_0^4 x^4 (4 - x^2)\, dx = \frac{16}{7}$$

$$\mu_2' = \int_{-\infty}^{\infty} x^2\, f(x)\, dx = k\int_0^4 x^5\, (4-x)^2\, dx = \frac{40}{7}$$

$$\text{Variance} = \mu_2 = \mu_2' - \mu_1'^2 = \left\{\frac{40}{7} - \left(\frac{16}{7}\right)^2\right\} = \frac{24}{49}$$

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

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

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

Let the events A and B be defined as follows:

A: One waits between 0 to 2 minutes inclusive.

B: One waits between 0 to 3 minutes inclusive.

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

Sol. (*a*)

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

Scanned with CamScanner

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

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

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

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

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

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

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

$$\therefore \quad P(\overline{A} \cap \overline{B}) = P(X > 3) = \int_3^\infty f(x)\,dx = \int_3^6 f(x)\,dx + \int_6^\infty f(x)\,dx$$

$$= \int_3^6 \frac{1}{9}\,dx = \frac{1}{9}\,|\,x\,|_3^6 = \frac{1}{3}$$

TEST YOUR KNOWLEDGE

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

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

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

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

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

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

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

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

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

6. What do you understand by probability distribution?

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

Scanned with CamScanner

8. Show that the symmetrical distribution

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

represents a probability density function. Also show that

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

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

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

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

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

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

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

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

Answers

1. $b - a$ **2.** $y_0 = \frac{1}{\sigma}$ **3.** $\frac{3}{16}$ **4.** $\frac{4}{9}$

7. $k = \frac{1}{2}$, mean = 3, variance = 3, $\beta_1 = \frac{4}{3}$, $\beta_2 = 5$

9. $\frac{1}{25}$ **10.** (*i*) 0.3935 (*ii*) 0.1481 (*iii*) 0.2231

4.26 JOINT PROBABILITY DISTRIBUTION

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

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

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

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

(*i*) $p(x_i, y_j) \ge 0$ (*ii*) $\sum_i \sum_j p(x_i, y_j) = 1.$

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

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

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

Scanned with CamScanner

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

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

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

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

4.27 MARGINAL PROBABILITY DISTRIBUTION

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

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

and

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

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

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

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

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

and similarly for fixed y_j

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

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

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

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

Let X and Y be continuous random variables.

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

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

and marginal probability density function of Y is

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

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

4.28 CONDITIONAL PROBABILITY DISTRIBUTION

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

Scanned with CamScanner

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

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

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

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

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

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

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

ILLUSTRUTIVE EXAMPLES

Example 1. *The joint probability distribution of two random variables X and Y is given by:* $P(X = 0, Y = 1) = \frac{1}{3}$, $P(X = 1, Y = -1) = \frac{1}{3}$, *and* $P(X = 1, Y = 1) = \frac{1}{3}$.

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

Sol. (*i*) $P(X = -1) = \sum_y P(X = -1, Y = y)$

$$= P(X = -1, Y = -1) + P(X = -1, Y = 0) + P(X = -1, Y = 1) = 0$$

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

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

Scanned with CamScanner

Thus

Marginal distribution of X is:

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

Marginal distribution of Y is:

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

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

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

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

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

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

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

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

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

(i) $P(X \le 1, Y = 2)$, *(ii)* $P(X \le 1)$,

(iii) $P(Y \le 3)$, *and* *(iv)* $P(X < 3, Y \le 4)$.

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

Sol. The marginal distributions are given below:

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

Scanned with CamScanner

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

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

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

(iv) $P(X < 3, Y \le 4) = P(X = 0, Y \le 4) + P(X = 1, Y \le 4) + P(X = 2, Y \le 4)$

$$= \left(\frac{1}{32} + \frac{2}{32}\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{32} + \frac{1}{32} + \frac{1}{64} + \frac{1}{64}\right) = \frac{9}{16}.$$

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

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

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

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

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

$$\therefore \quad P(X = 1) = \sum_y P(X = 1, Y = y)$$

$$= P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 1, Y = 4)$$

$$= \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36}.$$

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

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

Similarly, we can obtain the marginal distribution of Y.

Marginal Distribution of X

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

Marginal Distribution of Y

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

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

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

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

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

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

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

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

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

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

y :	1	2	3	4
$P(Y = y \mid X = 2)$:	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{9}$

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

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

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

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

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

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

$$p_X(x) = \sum_{y=0}^{x} p(x, y) = \sum_{y=0}^{x} \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} = \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^{x} \frac{x!\, p^y (1-p)^{x-y}}{y!(x-y)!}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} \sum_{y=0}^{x} {}^x C_y p^y (1-p)^{x-y} = \frac{\lambda^x e^{-\lambda}}{x!} [p + (1-p)]^x = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

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

$$p_Y(y) = \sum_{x=y}^{\infty} p(x, y) = \sum_{x=y}^{\infty} \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} = \frac{(\lambda p)^y e^{-\lambda}}{y!} \sum_{x=y}^{\infty} \frac{[\lambda(1-p)]^{x-y}}{(x-y)!}$$

$$[\because\ y = 0, 1, 2, \ldots, x \Rightarrow x \le y \Rightarrow x \ge y]$$

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

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

Scanned with CamScanner

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

$$P_{Y\mid X}(y \mid x) = \frac{p_{XY}(x, y)}{p_X(x)} = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y} x!}{y!(x-y)!\lambda^x e^{-\lambda}} = \frac{x!}{y!(x-y)!} p^y (1-p)^{x-y}$$

$$= {}^xC_y p^y (1-p)^{x-y}, x \geq y \quad i.e., \quad y = 0, 1, 2, ..., x.$$

The conditional probability distribution of X for given Y is:

$$p_{X\mid Y}(x \mid y) = \frac{p_{XY}(x, y)}{p_Y(y)} = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{y!(x-y)!} \cdot \frac{y!}{e^{-\lambda p}(\lambda p)^y} \qquad [c.f. \text{ Part } (i)]$$

$$= \frac{e^{-\lambda q}(\lambda q)^{x-y}}{(x-y)!}; \; q = 1-p, x \geq y \; i.e., x = y, y+1, y+2, ...$$

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

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

Find (i) P (X < 1 ∩ Y < 3), (ii) P (X + Y < 3), and (iii) P (X < 1 | Y < 3).

Sol. We have

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

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

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

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

Example 6. *Suppose that two-dimensional continuous random variable (X, Y) has joint p.d.f. given by*

$$f(x, y) = \begin{cases} 6x^2 y, 0 < x < 1, 0 < y < 1 \\ 0, \text{ elsewhere} \end{cases}$$

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

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

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

(ii) $P\left(0 < X < \frac{3}{4}, \frac{1}{3} < Y < 2\right) = \int_0^{3/4} \int_{1/3}^1 6x^2 y\, dxdy + \int_0^{3/4} \int_1^2 0 \,.\, dxdy$

$$= \int_0^{3/4} 6x^2 \left|\frac{y^2}{2}\right|_{1/3}^1 dx = \frac{8}{9} \int_0^{3/4} 3x^2\, dx = \frac{8}{9} \left| x^3 \right|_0^{3/4} = \frac{3}{8}.$$

Scanned with CamScanner

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

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

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

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

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

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

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

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

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

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

(iii) Check for independence of X and Y.

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

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

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

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

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

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

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

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

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

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

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

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

Scanned with CamScanner

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

$$f_X(x) = \int_0^\infty f(x, y)\, dy = \frac{9}{2(1+x)^4} \int_0^\infty \frac{(1+y)+x}{(1+y)^4}\, dy$$

$$= \frac{9}{2(1+x)^4} \int_0^\infty \{(1+y)^{-3} + x(1+y)^{-4}\}\, dy$$

$$= \frac{9}{2(1+x)^4} \left(\left| \frac{-1}{2(1+y)^2} \right|_0^\infty + x \left| \frac{-1}{3(1+y)^3} \right|_0^\infty \right)$$

$$= \frac{9}{2(1+x)^4} \cdot \left(\frac{1}{2} + \frac{x}{3} \right) = \frac{3}{4} \cdot \frac{3+2x}{(1+x)^4};\ 0 < x < \infty$$

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

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

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

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

Example 9. *Joint distribution of X and Y is given by:* $f(x, y) = 4xy\, e^{-(x^2+y^2)}$; $x \ge 0$, $y \ge 0$. *Test whether X and Y are independent. For the above joint distribution, find the conditional density of X given Y = y.*

Sol. Marginal density of X is given by

$$f_X(x) = \int_0^\infty f_{XY}(x, y)\, dy = \int_0^\infty 4xy\, e^{-(x^2+y^2)}\, dy = 4x\, e^{-x^2} \int_0^\infty y\, e^{-y^2}\, dy$$

$$= 4x\, e^{-x^2} \cdot \int_0^\infty e^{-t} \cdot \frac{dt}{2} = 2x \cdot e^{-x^2} \left| -e^{-t} \right|_0^\infty$$

$$\therefore \quad f_X(x) = 2x\, e^{-x^2};\ x \ge 0$$

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

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

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

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

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

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

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

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

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

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

$$P(X^2 \le x) = P(|X| \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_X(x)\,dx = \sqrt{x} \qquad ...(1)$$

$$P(X^2 \le x \cap Y^2 \le y) = P(|X| \le \sqrt{x} \cap |Y| \le \sqrt{y})$$

$$= \int_{-\sqrt{x}}^{\sqrt{x}} \left[\int_{-\sqrt{y}}^{\sqrt{y}} f(u, v)\,dv \right] du = \sqrt{x}\,\sqrt{y}$$

$$= P(X^2 \le x)\,.\,P(Y^2 \le y) \qquad \text{[From (1)]}$$

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

TEST YOUR KNOWLEDGE

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

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

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

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

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

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

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

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

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

   (*i*) Verify that the area under the curve is unity.
   (*ii*) Find the mean and variance of the above distribution.

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

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

   (*a*) Find the value of A that makes $f(x)$ a p.d.f.
   (*b*) What is the probability that the number of minutes that she will talk over the phone is: (*i*) more than 10 minutes, (*ii*) less than 5 minutes, and (*iii*) between 5 and 10 minutes.

Scanned with CamScanner

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

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

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

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

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

(i) Determine the constant k.

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

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

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

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

Find the marginal and conditional distributions.

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

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

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

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

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

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

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

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

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

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

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

(ii) Are X and Y independent?

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

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

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

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

Scanned with CamScanner

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

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

Find the conditional distribution of Y for $X = x$.

Answers

1. Mean = 4, Variance = 80
2. Mean = b, Variance = b^2
3. A = 1, B = 0
4. (*i*) $c = 12$, (*ii*) Mean = 3/5
5. (*ii*) Mean = 0, Variance = 1
6. (*a*) $A = \frac{1}{5}$ (*b*) (*i*) $\frac{1}{e^2}$ (*ii*) $\frac{e-1}{e}$ (*iii*) $\frac{e-1}{e^2}$
7. (*a*) $k = 8/9$ (*b*) (*i*) $1 - \frac{1}{e}$ (*ii*) $\frac{1}{2}$ (*iii*) $1 - \frac{2}{e}$
8. (*i*) $k = 1$ (*ii*) $f_1(x \mid y) = e^{-x}$ (*iii*) e^{-3}
9. (*i*) $f_X(x) = \begin{cases} 4x(1-x^2), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$, $f_Y(y) = 4y^3, 0 < y < 1$

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

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

10. (*a*) (*i*) $f_X(x) = \begin{cases} \frac{4}{9}x(4-x^2), & 1 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$ (*ii*) $f_{X|Y}(x \mid y) = \frac{2x}{y^2 - 1}, \quad 1 \le x \le y$

    $f_Y(y) = \frac{4}{9}y(y^2 - 1), 1 \le y \le 2$ $f_{Y|X}(y \mid x) = \frac{2y}{4-x^2}, x \le y \le 2.$

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

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

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

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

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

13.

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

Scanned with CamScanner

14. Conditional distribution of Y for X = x.

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

4.29 MATHEMATICAL EXPECTATION OR EXPECTED VALUE OF A RANDOM VARIABLE

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

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

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

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

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

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

$$E(x^r) = p_1x_1^{\,r} + p_2x_2^{\,r} + ... + p_nx_n^{\,r}.$$

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

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

In particular $\mu_1' = E(x) = p_1x_1 + p_2x_2 + ... = \sum_i p_ix_i$

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

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

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

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

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

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

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

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

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

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

is expectation of continuous random variable with certain restrictions.

4.30 LAWS OF EXPECTATION

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

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

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

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

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

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

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

Then $$E(aX) = ax_1p_1 + ax_2p_2 + \ldots \qquad \{\because\ P(aX = ax_i) = P(X = x_i)\}$$

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

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

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

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

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

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

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

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

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

Proof. $$E(aX + b) = E(aX) + E(b), \qquad \{\because\ E(X + Y) = E(X) + E(Y)\}$$

$$= aE(X) + b. \qquad \{\because\ E(aX) = aE(X)\}$$

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

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

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

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

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

Scanned with CamScanner

Theorem 5. *Expectation of the difference of two random variables.*

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

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

In general

$$E(X_1 \pm X_2 \pm, \ldots, \pm X_k) = E(X_1) \pm E(X_2) \pm \ldots \pm E(X_k).$$

ILLUSTRITIVE EXAMPLES

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

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

Hence, $$E(x) = p_1x_1 + p_2x_2 + \ldots + p_6x_6 = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \cdots + \frac{1}{6} \cdot 6 = 3.5$$

Also, $$\text{var}(x) = E(x^2) - [E(x)]^2 = \frac{1}{6}(1^2 + 2^2 + \ldots + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Example 2. *Thirteen cards are drawn simultaneously from a deck of 52. If aces count 1, face cards 10 and others according to denomination, find the expectation of the total score on the 13 cards.*

Sol. Let x_i be the number corresponding to the i^{th} card, then x_i takes the values 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 10, 10; 10 each having the probability $\frac{1}{13}$. Hence

$$E(x_i) = \frac{1}{13} \cdot 1 + \frac{1}{13} \cdot 2 + \frac{1}{13} \cdot 3 + \cdots + \frac{1}{13} \cdot 9 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10 + \frac{1}{13} \cdot 10$$

$$E(x_i) = \frac{1}{13}(1 + 2 + 3 + \ldots + 9 + 10 + 10 + 10 + 10) = \frac{85}{13}.$$

Example 3. *In four tosses of a coin, let x be the number of heads. Calculate the expected values of x.*

Sol. There will be all the heads, there is only one way

i.e., $$P(x = 4) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$$

$$P(x = 3) = \frac{{}^4C_3}{16} = \frac{4}{16} = \frac{1}{4}$$

$$P(x = 2) = \frac{{}^4C_2}{16} = \frac{6}{16} = \frac{3}{8}$$

$$P(x = 1) = \frac{{}^4C_1}{16} = \frac{4}{16} = \frac{1}{4}$$

$$P(x = 0) = \frac{{}^4C_0}{16} = \frac{1}{16}.$$

Scanned with CamScanner

Hence the probability distribution of x is

x:	0	1	2	3	4
P(x):	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

$$\therefore \quad E(x) = \sum_{x=0}^{4} x(px) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16} = 2.$$

Example 4. *Find $E(x)$, $E(x^2)$, $E\{(x - \bar{x})^2\}$ for the following probability distribution:*

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

$$E(x) = \Sigma x \,.\, p(x) = 8 \cdot \frac{1}{8} + 12 \cdot \frac{1}{6} + 16 \cdot \frac{3}{8} + 20 \cdot \frac{1}{4} + 24 \cdot \frac{1}{12} = 16.$$

This represents the *mean* of the distribution.

$$E(x^2) = \Sigma x^2 \,.\, p(x) = 8^2 \cdot \frac{1}{8} + 12^2 \cdot \frac{1}{6} + 16^2 \cdot \frac{3}{8} + 20^2 \cdot \frac{1}{4} + 24^2 \cdot \frac{1}{12} = 276.$$

This represents the *second moment* about the origin zero.

$$E\{(x - \bar{x})^2\} = \Sigma(x - \bar{x})^2\, p(x)$$

$$= (8-16)^2 \cdot \frac{1}{8} + (12-16)^2 \cdot \frac{1}{6} + (16-46)^2 \cdot \frac{3}{8} + (20-16)^2 \cdot \frac{1}{4} + (24-16)^2 \cdot \frac{1}{12} = 20$$

This represents the *variance of the distribution.*

4.31 INDEPENDENT VARIATES

The variates, x and y, are said to be independent, if the probability that either of them will assume a prescribed value does not depend on the value assumed by the other.

3.31.1 Theorem. *The expectation of the product of two independent variates is equal to the product of their expectations i.e., $E(xy) = E(x)E(y)$.*

4.32 COVARIANCE

If x and y are two variates with the respective expected values (or means) x and y, the covariance between x and y is defined as

$$\text{cov}\,(x, y) = E[(x - \bar{x})\,(y - \bar{y})].$$

Thus the expected value of product of the deviations of the two variates from their mean is called their *covariance*.

4.33 THEOREM

The covariance of two independent variances is equal to zero.

If x and y are independent variates, then

$$E(x - \bar{x}) = E(x) - E(\bar{x}) = \bar{x} - \bar{x} = 0,$$

$$E(y - \bar{y}) = E(y) - E(\bar{y}) = \bar{y} - \bar{y} = 0.$$

$$\therefore \quad \text{cov}\,(x, y) = E(x - \bar{x})E(y - \bar{y}) = 0$$

Scanned with CamScanner

Note.

$$\begin{aligned}\text{cov}(x, y) &= E\{(x - \bar{x})(y - \bar{y})\} = E(xy - \bar{x}y - \bar{y}x + \bar{x}\bar{y})\\ &= E(xy) - E(\bar{x}y) - E(\bar{y}x) + E(\bar{x}\bar{y})\\ &= E(xy) - \bar{x}E(y) - \bar{y}E(x) + \bar{x}\bar{y}\\ &= E(xy) - \bar{x}\bar{y} - \bar{y}\bar{x} + \bar{x}\bar{y}\\ &= E(xy) - E(x)E(y).\end{aligned}$$

4.34 VARIANCE OF *n* VARIATES

If $x_1, x_2, \ldots, x_n$ be n variates with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ to find the variance of u where

$$u = a_1x_1 + a_2x_2 + \ldots + a_nx_n \text{ where } a\text{'s are constant.}$$

$$\begin{aligned}E(u) &= E(a_1x_1 + a_2x_2 + \ldots + a_nx_n)\\ &= E(a_1x_1) + E(a_2x_2) + \ldots + E(a_nx_n)\\ &= a_1E(x_1) + a_2E(x_2) + \ldots + a_nE(x_n)\end{aligned}$$

$$\therefore \quad u - E(u) = a_1[x_1 - E(x_1)] + a_2[x_2 - E(x_2)] + \cdots + a_n[x_n - E(x_n)]$$

Squaring both sides, we get

$$\begin{aligned}[u - E(u)]^2 &= a_1^2[x_1 - E(x_1)]^2 + a_2^2[x_2 - E(x_2)]^2 + \ldots\\ &\quad + a_n^2[x_n - E(x_n)]^2 + 2a_1a_2[x_1 - E(x_1)][x_2 - E(x_2)] + \ldots\end{aligned}$$

Taking expected values of both sides, we get

$$\begin{aligned}\text{var}(u) &= a_1^2 \text{ var}(x_1) + a_2^2 \text{ var}(x_2) + \ldots + a_n^2 \text{ var}(x_n)\\ &\quad + 2a_1a_2 \text{ cov}(x_1, x_2) + \ldots + 2a_{n-1}a_n \text{ cov}(x_{n-1}, x_n)\end{aligned}$$

Cor. 1. If $a_1 = a_2 = 1$, $a_3 = a_4 = \ldots = a_n = 0$, we have

$$\text{var}(x_1 + x_2) = \text{var}(x_1) + \text{var}(x_2) + 2\text{ cov}(x_1, x_2).$$

Cor. 2. If $a_1 = 1$, $a_2 = -1$, $a_3 = a_4 = \ldots = a_n = 0$, we have

$$\text{var}(x_1 - x_2) = \text{var}(x_1) + \text{var}(x_2) - 2\text{ cov}(x_1, x_2).$$

Cor. 3. If $x_1, x_2, \ldots x_n$ are independent, we have

$$\text{var}(u) = a_1^2 \text{ var}(x_1) + a_2^2 \text{ var}(x_2) + \cdots + a_n^2 \text{ var}(x_n).$$

ILLUSTRATIVE EXAMPLES

Example 1. *Balls are taken one by one out of an urn containing a white and b black balls until the first white ball is drawn. Show that the expectation of the number of black balls preceding the first white ball is* $\dfrac{b}{a+1}$.

Sol. Let x_1 denote the number of black balls before the first white ball is drawn and x_i denote the number of black balls following the $(i-1)^{\text{th}}$ white balls and preceding the i^{th} white balls.

Let x_{a+1} denote the number of black balls left after the white balls have been drawn.

Then $$x_1 + x_2 + \cdots + x_{a+1} = b,$$

Hence $$E(x_1 + x_2 + \cdots + x_{a+1}) = E(b),$$

or $$E(x_1) + E(x_2) + \cdots + E(x_{a+1}) = E(b).$$

The probability of every system of numbers $x_1, x_2, \ldots, x_{a+1}$ is the same and is $\dfrac{a!\,b!}{(a+b)!}$

Scanned with CamScanner

Thus, $E(x_1) = E(x_2) = \ldots = E(x_{a+1}) = K$ say.

Therefore, $(a+1)K = b$ giving $K = \dfrac{b}{a+1}$.

Example 2. *A box contains a white and b black balls ; c balls are drawn. Show that the expectation of the number of white balls drawn is* $\dfrac{ca}{a+b}$.

Sol. Let a variate x_i be defined as follows

$$x_i = x_1, x_2, x_3, \ldots, x_c$$

and in that case

$$x_i = \begin{cases} 1 \text{ if the } i^{\text{th}} \text{ ball drawn is white} \\ 0 \text{ if the } i^{\text{th}} \text{ ball drawn is black} \end{cases}$$

Then the number of white balls, s, among the c drawn balls is given by

$$s = x_1 + x_2 + \cdots + x_c.$$

Now, $$P(x_i = 1) = \frac{a}{a+b},\ P(x_i = 0) = \frac{b}{a+b}$$

$\therefore$ $$E_i = 1 \cdot \frac{a}{a+b} + 0 \cdot \frac{b}{a+b} = \frac{a}{a+b}$$

$\therefore$ $$E(S) = \frac{a}{a+b} + \frac{a}{a+b} + \cdots + \frac{a}{a+b},\ c \text{ terms} = \frac{ca}{a+b}$$

Example 3. *What is the expectation of the number of failures preceding the first success in an infinite series of independent trials with constant probability p of success in a trial?*

Sol. Let x denote the number of failures preceding the first success we find that x takes the values 0, 1, 2, 3, 4, ... Their respective probabilities are $p, (1-p)p, (1-p)^2p, (1-p)^3p, \ldots$

Since, $1 - p = q$, these are $p, qp, q^2p, q^3p, \ldots$

$$E(x) = 0 \cdot p + 1 \cdot qp + 2q^2p + 3q^3p + \ldots = pq(1 + 2q + 3q^2 + \ldots)$$

$$= pq(1-q)^{-2} = pq\,\frac{1}{p^2} = \frac{q}{p} = \frac{1-p}{p} = \frac{1}{p} - 1.$$

Example 4. *Show by an example that the mathematical expectation need not be finite.*

Sol. Let x be the discontinuous random variable having values $0!, 1!, 2!, 3!, \ldots, \infty$ and let the probability law for it be $\dfrac{e^{-1}}{x!}$

$$E(x!) = 0!\,.\frac{e^{-1}}{0!} + 1!\,.\frac{e^{-1}}{1!} + 2!\,.\frac{e^{-1}}{2!} + 3!\,.\frac{e^{-1}}{3!} + \ldots\ldots$$

$$= e^{-1} + e^{-1} + e^{-1} + e^{-1} + \ldots\ldots$$

$$= e^{-1}(1 + 1 + \ldots\ldots \infty)$$

$$= e^{-1} \sum_{0}^{\infty} 1, \text{ which is not finite.}$$

Example 5. *Show that if x takes the values $x_n = (-1)^n 2^n n^{-1}$ for n = 1, 2, ... with probabilities $p_n = 2^{-n}$ then $E(x) = -\log 2$.*

Sol. Here

$$E(x) = \sum_{n=1}^{\infty} 2^{-n} \cdot (-1)^n \cdot 2^n n^{-1}$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \ldots$$

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

Example 6. *What is the mathematical expectation of the sum of points on n dice?*

Sol. Denoting by x_i the number of points on the i^{th} dice, the sum of points on n dice is

$$x_1 + x_2 + \cdots + x_n$$

$$\therefore \quad E(x_1 + x_2 + \cdots + x_n) = E(x_1) + E(x_2) + \cdots + E(x_n)$$

But for every single dice $E(x_i) = \frac{7}{2}$, $i = 1, 2, 3, \ldots, n$

$$\therefore \quad E(x_1 + x_2 + \cdots + x_n) = \frac{7}{2} + \frac{7}{2} + \cdots n \text{ terms} = \frac{7}{2} n.$$

Example 7. *Find the excepted value of the product of points on n dice.*

Sol. Denoting by x_i the number of points on the i^{th} dice the product of points on n dice is

$$x_1, x_2, x_3, \ldots, x_n$$

For every single dice $E(x_i) = \frac{7}{2}$, $i = 1, 2, 3, \ldots, n$

$$\therefore \quad E(x_1 x_2 \ldots x_n) = E(x_1) \cdot E(x_2) \ldots E(x_n) = \frac{7}{2} \cdot \frac{7}{2} \cdot \frac{7}{2} \ldots n \text{ factors} = \left(\frac{7}{2}\right)^n.$$

Example 8. *There is a series of n independent trials, where p_i is the probability for i^{th} trial. Show that the mean number of successes is $\sum_{i=1}^{n} p_i$.*

Sol. Let a variable x_i be associated with i^{th} trial, which has the value 1 in case of success and the value 0 in case of a failure. If m be the number of successes in n trials then

$$m = x_1 + x_2 + \ldots + x_n$$

$$\therefore \quad E(m) = E(x_1 + x_2 + \ldots + x_n) = E(x_1) + E(x_2) + \ldots + E(x_n)$$

Since x_i can take the values 1 or 0 with probabilities p_i or $1 - p_i$, we have

$$E(x_i) = 1, p_i + 0\,(1 - p_i) = p_i$$

$$\therefore \quad E(m) = E(x_1) + E(x_2) + \ldots + E(x_n) = p_1 + p_2 + \ldots + p_n.$$

Example 9. *Find the variance of the number of successes in a series of n independent trials in which the probability of success in the i^{th} trial is p_i.*

Sol. Here

$$m = x_1 + x_2 + \ldots + x_n$$

$$\therefore \quad \text{var}(m) = \text{var}(x_1 + x_2 + \ldots + x_n) = \text{var}(x_1) + \text{var}(x_2) + \ldots + \text{var}(x_n)$$

If there is success, $x_i^2 = 1$, if failure, $x_i^2 = 0$

$$\therefore \quad E(x_i^2) = 1 \cdot p_i + 0 \cdot (1 - p_i) = p_i.$$

Scanned with CamScanner

Here the variance of the number of successes in the i^{th} trial is given by

$$\begin{aligned}\text{var}(x_i) &= E[x_i - E(x_i)]^2\\ &= E(x_i)^2 - 2x_i\, E(x_i) + \{E(x_i)\}^2 = E(x_i)^2 - \{E(x_i)\}^2\\ &= p_i - p_i^2 = p_i(1-p_i) = p_i\, q_i\end{aligned}$$

$$\therefore \qquad \text{var}(m) = \sum_{i=1}^{n} p_i q_i = p_1q_1 + p_2q_2 + \dots + p_nq_n.$$

Corollary 1. Show that var $(m) \le \frac{1}{4}$.

Since $pq = p(1-p) = p - p^2 - \frac{1}{4} + \frac{1}{4} = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \le \frac{1}{4}$, or $\left(p - \frac{1}{2}\right)^2$ is positive.

Corollary 2. If $p_1 = p_2 = \dots = p_n = p$, then var $(m) = pq + pq + \dots + pq = npq$.

Example 10. *From a point on the circumference of a circle of radius a, chord is drawn in a random direction. Show that the expected value of the length of the chord is $4a/\pi$ and that the variance of the length is $2a^2(1 - 8/\pi^2)$. Also show that the chance is $\frac{1}{3}$ and the length of the chord will exceed the length of the side of an equilateral triangle inscribed in the circle.*

Sol. Let C be the centre and A be any point. Let AB be any random chord.

Let $\angle DAB = \theta$. Then θ lies from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

$$dF(\theta) = f(\theta)\, d\theta = \frac{d\theta}{\frac{1}{2}\pi - \left(-\frac{1}{2}\pi\right)} = \frac{d\theta}{\pi}, \quad -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$$

If a be the radius, AB $= 2a \cos\theta$

$$\therefore \qquad E(AB) = \int_{-\pi/2}^{\pi/2} AB\,.\, f(\theta) = \frac{2a}{\pi}\int_{-\pi/2}^{\pi/2} \cos\theta\, d\theta = \frac{4a}{\pi}.$$

$$E[(AB)^2] = \int_{-\pi/2}^{\pi/2} (AB)^2\, f(\theta) d\theta = \frac{4a^2}{\pi}\int_{-\pi/2}^{\pi/2} \cos^2\theta\, d\theta = 2a^2.$$

$$\text{var}(AB) = E[(AB)^2] - \{E(AB)\}^2 = 2a^2 - 16a^2/\pi^2 = 2a^2(1 - 8/\pi^3).$$

Now as the inscribed equilateral triangle is $a\sqrt{3}$, therefore

$$\begin{aligned}P(AB > a\sqrt{3}) &= P(2a\cos\theta > a\sqrt{3}) = P(\cos\theta > \sqrt{3}/2)\\ &= P\left(|\theta| < \frac{1}{6}\pi\right) = P\left(-\frac{1}{6}\pi < \theta < \frac{1}{6}\pi\right)\\ &= \int_{-\pi/6}^{\pi/2} f|\theta|\, d\theta = \frac{1}{\pi}\int_{-\pi/6}^{\pi/6} 1\,.\, d\theta = \frac{1}{\pi}\cdot\frac{\pi}{3} = \frac{1}{3}.\end{aligned}$$

Scanned with CamScanner

TEST YOUR KNOWLEDGE

If x is a variate and a is a constant, show that:

1. $E(a) = a$.
2. $E(ax) = aE(x)$.
3. $\text{var}(ax + b) = a^2 . \text{var}(x)$.
4. $\text{cov}(ax, by) = ab \text{ cov}(x, y)$.
5. $\text{cov}(x + a, y - b) = \text{cov}(x, y)$.
6. $\text{cov}(x, x) = \text{var}(x)$.
7. Prove that:
$$\text{cov}(ax + by, cx + dy) = ac \text{ var}(x) + bd \text{ var}(y) + (ad + bc) \text{ cov}(x, y).$$
8. If x, y, z are random variables each with expectation 10 and variance 1, 4 and 9 respectively; $r(x, y) = 0, r(y, z) = \frac{1}{4}, r(x, z) = \frac{1}{4}$ obtain

   (a) $E(x + y - 2z)$
   (b) $\text{cov}(x + z, y + z)$
   (c) $V(x - 2z)$
   (d) $\text{cov}(3x, 5z)$.
9. Two unbiased dice are thrown. Find the expected value of the sum of numbers of points drawn.
10. A coin is tossed until a head appears. What is the expectation of the number of tosses?
11. In a lottery m tickets are drawn out of n tickets numbered from 1 to n. What is the expectation and the variance of the sum of the numbers drawn?
12. Three urns contain respectively 3 green and 2 white balls, 5 green and 6 white balls and 2 green and 4 white balls. One ball is drawn from each urn. Find the expected number of white balls drawn out.

$$\left[\textbf{Hint. } 0\left(\frac{3}{5}.\frac{5}{11}.\frac{2}{6}\right) + 1\cdot\left(\frac{2}{5} + \frac{2}{5}.\frac{2}{6} + \frac{3}{5}.\frac{5}{11}\frac{4}{6} + \frac{3}{5}.\frac{6}{11}.\frac{2}{6}\right) + 2\left(\frac{2}{5}.\frac{6}{11}\frac{2}{6} + \frac{2}{5}.\frac{5}{11}\frac{4}{6} + \frac{3}{5}.\frac{6}{11}.\frac{4}{6}\right) + 3\cdot\left(\frac{2}{5}.\frac{6}{11}.\frac{4}{6}\right) = \frac{266}{165}\right]$$

13. Show that the expectation of the number of failures preceding the first success in an indefinite series of independent trials with constant probability p of success, is
$$qp + 2q^2p + 3q^3p + \cdots = \frac{pq}{(1-q)^2} = \frac{q}{p} = \frac{1-p}{p}.$$
14. Define mathematical expectation. Prove that mathematical expectation of the sum of two discrete random variables is equal to the sum of their mathematical expectations.
15. x denotes the profit that a man can make in business. He may earn ₹2,800 with probability $\frac{1}{2}$, he may lose ₹5,000 with probability $\frac{3}{10}$ and he may neither lose nor gain with probability $\frac{1}{5}$. Show that the mathematical expectation is – 100.
16. If x is a random variable for which $E(x) = 10$ and $\text{var}(x) = 25$. Find the positive values of a and b such that $y = ax - b$ has expectation 0 and variance 1.

Answers

9. 7 **10.** 2 **12.** $\frac{266}{165}$

11. $\frac{1}{2}m(n+1); \frac{1}{12}m(n+1)(n-2)$ **16.** $a = \frac{1}{5}, b = 2$.

Scanned with CamScanner

4.35 THEORETICAL PROBABILITY DISTRIBUTIONS

Generally, frequency distribution are formed from the observed or experimental data. However, frequency distribution of certain populations can be deduced mathematically by fitting theoretical probability distribution under certain assumptions.

Frequency distributions can be classified under two heads:

(*i*) Observed Frequency Distributions.

(*ii*) Theoretical or Expected Frequency Distributions.

Observed frequency distributions are based on actual observation and experimentation. If certain hypothesis is assumed, it is sometimes possible to derive mathematically what the frequency distribution of certain universe should be. Such distributions are called **Theoretical Distributions.**

Theoretical probability distributions are of two types:

(*i*) Discrete probability distribution. Binomial, poisson, geometric, negative binomial, hypergeometric, multinomial, multivariate hypergeometric distributions.

(*ii*) Continuous probability distributions.

Uniform, normal Gamma, exponential, χ^2, Beta, bivariate normal, t, F-distributions.

Here, we will study three important theoretical probability distributions:

1. Binomial Distribution (or Bernoulli's Distribution)
2. Poisson's Distribution
3. Normal Distribution.

4.36 BINOMIAL PROBABILITY DISTRIBUTION (*A.K.T.U. 2018*)

It was discovered by a Swiss Mathematician Jacob James Bernoulli in the year 1700.

This distribution is concerned with trials of a repretitive nature in which only the occurrence or non-occurrence, success or failure, acceptance or rejection, yes or no of a particular event is of interest.

For convenience, we shall call the occurrence of the event 'a success' and its non-occurrence 'a failure'.

Let there be n independent trials in an experiment. Let a random variable X denote the number of successes in these n trials. Let p be the probability of a success and q that of a failure in a single trial so that $p + q = 1$. Let the trials be independent and p be constant for every trial.

Let us find the probability of r successes in n trials.

r successes can be obtained in n trials in ${}^nC_r$ ways.

$$\therefore \quad P(X = r) = {}^nC_r P(\underbrace{S\,S\,S \ldots\ldots S}_{r \text{ times}} \quad \underbrace{F\,F\,F \ldots\ldots F}_{(n-r) \text{ times}})$$

$$= {}^nC_r \underbrace{P(S)\,P(S) \ldots\ldots P(S)}_{r \text{ factors}} \quad \underbrace{P(F)\,P(F) \ldots\ldots P(F)}_{(n-r) \text{ factors}}$$

$$= {}^nC_r \underbrace{p\,p\,p \ldots\ldots p}_{r \text{ factors}} \quad \underbrace{q\,q\,q \ldots\ldots q}_{(n-r) \text{ factors}}$$

$$= {}^nC_r p^r q^{n-r} \qquad \ldots(1)$$

Scanned with CamScanner

Hence $\boxed{P(X = r) = {}^nC_r\, p^r q^{n-r}}$ where $p + q = 1$ and $r = 0, 1, 2, ..., n$.

The distribution (1) is called the *binomial probability distribution* and X is called the *binomial variate*.

Note 1. $P(X = r)$ is usually written as $P(r)$.

Note 2. The successive probabilities $P(r)$ in (1) for $r = 0, 1, 2, \ldots\ldots, n$ are

$${}^nC_0 q^n, {}^nC_1 q^{n-1} p, {}^nC_2 q^{n-2} p^2, \ldots\ldots, {}^nC_n p^n$$

which are the successive terms of the binomial expansion of $(q + p)^n$. That is why this distribution is called "binomial" distribution.

Note 3. n and p occurring in the binomial distribution are called the *parameters* of the distribution.

Note 4. In a binomial distribution:

(*i*) n, the number of trials is finite.

(*ii*) each trial has only two possible outcomes usually called success and failure.

(*iii*) all the trials are independent.

(*iv*) p (and hence q) is constant for all the trials.

4.37 RECURRENCE OR RECURSION FORMULA FOR THE BINOMIAL DISTRIBUTION

In a binomial distribution,

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

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

$$\therefore \quad \frac{P(r+1)}{P(r)} = \frac{(n-r)!}{(n-r-1)!} \times \frac{r!}{(r+1)!} \times \frac{p}{q}$$

$$= \frac{(n-r) \times (n-r-1)!}{(n-r-1)!} \times \frac{r!}{(r+1) \times r!} \times \frac{p}{q} = \left(\frac{n-r}{r+1}\right) \cdot \frac{p}{q}$$

$$\Rightarrow \quad \boxed{P(r+1) = \frac{n-r}{r+1} \cdot \frac{p}{q}\, P(r)}$$

which is the required recurrence formula. Applying this formula successively, we can find P(1), P(2), P(3), ..., if P(0) is known.

4.38 MEAN AND VARIANCE OF THE BINOMIAL DISTRIBUTION [A.K.T.U. 2018]

For the binomial distribution, $P(r) = {}^nC_r\, q^{n-r} p^r$

$$\text{Mean } \mu = \sum_{r=0}^{n} r P(r) = \sum_{r=0}^{n} r \,.\, {}^nC_r q^{n-r} p^r$$

$$= 0 + 1 \,.\, {}^nC_1\, q^{n-1} p + 2 \,.\, {}^nC_2\, q^{n-2} p^2 + 3 \,.\, {}^nC_3\, q^{n-3} p^3 + \cdots\cdots + n \,.\, {}^nC_n\, p^n$$

$$= nq^{n-1} p + 2 \,.\, \frac{n(n-1)}{2 \cdot 1}\, q^{n-2} p^2 + 3 \,.\, \frac{n(n-1)(n-2)}{3 \cdot 2 \cdot 1}\, q^{n-3} p^3 + \cdots\cdots + n.p^n$$

Scanned with CamScanner

$$= nq^{n-1}p + n(n-1)\,q^{n-2}p^2 + \frac{n(n-1)(n-2)}{2.1}\,q^{n-3}p^3 + \ldots\ldots + np^n$$

$$= np[{}^{n-1}C_0q^{n-1} + {}^{n-1}C_1q^{n-2}p + {}^{n-1}C_2q^{n-3}p^2 + \ldots + {}^{n-1}C_{n-1}p^{n-1}]$$

$$= np(q+p)^{n-1} = np \qquad (\because\ p + q = 1)$$

Hence mean of the binomial distribution is np.

$$\text{Variance } \sigma^2 = \sum_{r=0}^{n} r^2\mathrm{P}(r) - \mu^2 = \sum_{r=0}^{n} [r + r(r-1)]\,\mathrm{P}(r) - \mu^2$$

$$= \sum_{r=0}^{n} r\mathrm{P}(r) + \sum_{r=0}^{n} r(r-1)\,\mathrm{P}(r) - \mu^2 = \mu + \sum_{r=2}^{n} r(r-1)\,{}^nC_r q^{n-r}p^r - \mu^2$$

(since the contribution due to $r = 0$ and $r = 1$ is zero)

$$= \mu + [2\cdot 1\cdot {}^nC_2q^{n-2}p^2 + 3\cdot 2\cdot {}^nC_3q^{n-3}p^3 + \ldots + n(n-1)\,{}^nC_np^n] - \mu^2$$

$$= \mu + \left[2\cdot 1\cdot \frac{n(n-1)}{2.1}\,q^{n-2}p^2 + 3\cdot 2\cdot \frac{n(n-1)(n-2)}{3.2.1}\,q^{n-3}p^3 + \ldots + n(n-1)p^n\right] - \mu^2$$

$$= \mu + [n(n-1)q^{n-2}p^2 + n(n-1)(n-2)q^{n-3}p^3 + \ldots + n(n-1)p^n] - \mu^2$$

$$= \mu + n(n-1)\,p^2[q^{n-2} + (n-2)q^{n-3}p + \ldots + p^{n-2}] - \mu^2$$

$$= \mu + n\,(n-1)p^2\,[{}^{n-2}C_0q^{n-2} + {}^{n-2}C_1q^{n-3}p + \ldots + {}^{n-2}C_{n-2}p^{n-2}] - \mu^2$$

$$= \mu + n(n-1)p^2(q+p)^{n-2} - \mu^2 = \mu + n(n-1)p^2 - \mu^2 \qquad [\because\ q + p = 1]$$

$$= np + n(n-1)p^2 - n^2p^2 = np[1-p] = npq. \qquad [\because\ \mu = np]$$

Hence the variance of the binomial distribution is npq.

Standard deviation of the binomial distribution is $\sqrt{npq}$.

4.39 MOMENT GENERATING FUNCTION OF BINOMIAL DISTRIBUTION [*A.K.T.U. 2018*]

1. About origin

$$\mathrm{M}_x(t) = \mathrm{E}(e^{tx}) = \sum_{x=0}^{n} e^{tx}\ {}^nC_x\ p^x\ q^{n-x} = \sum_{x=0}^{n} {}^nC_x\ (pe^t)^x\ q^{n-x} = (q + pe^t)^n$$

2. About mean [*G.B.T.U. 2012, U.P.T.U. 2015*]

$$\mathrm{M}_{x-np}(t) = \mathrm{E}[e^{t(x-np)}]$$

$$= e^{-npt}\,\mathrm{E}(e^{tx}) = e^{-npt}\,\mathrm{M}_x(t) = e^{-npt}\,(q + pe^t)^n$$

$$= (qe^{-pt} + pe^{t-pt})^n = (qe^{-pt} + pe^{qt})^n \qquad |\because\ 1 - p = q$$

4.40 MOMENTS ABOUT MEAN OF BINOMIAL DISTRIBUTION

$$\mathrm{M}_{x-np}(t) = (qe^{-pt} + pe^{qt})^n$$

$$= \left[q\left(1 - pt + \frac{p^2t^2}{2!} - \frac{p^3t^3}{3!} + \ldots\right) + p\left(1 + qt + \frac{q^2t^2}{2!} + \frac{q^3t^3}{3!} + \ldots\right)\right]^n$$

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

Scanned with CamScanner

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

$$= \left[1 + {}^nC_1 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq\,(q-p) + \frac{t^4}{4!} pq\,(1-3pq) + \ldots\right\}\right.$$

$$\left. + {}^nC_2 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq\,(q-p) + \ldots\right\}^2 + \ldots\right]$$

Now,

$$\mu_2 = \text{coefficient of } \frac{t^2}{2!} = npq$$

$$\mu_3 = \text{coefficient of } \frac{t^3}{3!} = npq\,(q-p)$$

$$\mu_4 = \text{coefficient of } \frac{t^4}{4!} = npq\,(1-3\,pq) + 3n(n-1)\,p^2q^2$$

$$= 3n^2p^2q^2 + npq\,(1-6pq)$$

Hence,

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\therefore \qquad \gamma_1 = \frac{1-2p}{\sqrt{npq}}$$

and

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

$$\therefore \qquad \gamma_2 = \frac{1-6pq}{npq}$$

Note 1. $\gamma_1 = \dfrac{1-2p}{\sqrt{npq}}$ gives a **measure of skewness** of the binomial distribution. If $p < \frac{1}{2}$, skewness is positive, if $p > \frac{1}{2}$, skewness is negative and if $p = \frac{1}{2}$, it is zero.

$\beta_2 = 3 + \dfrac{1-6pq}{npq}$ gives a **measure of the kurtosis** of the binomial distribution.

Note 2. If n independent trials constitute one experiment and this experiment is repeated N times then the frequency of r successes is $N \cdot {}^nC_r\, p^r\, q^{n-r}$.

4.41 APPLICATIONS OF BINOMIAL DISTRIBUTION

1. In problems concerning no. of defectives in a sample production line.
2. In estimation of reliability of systems.

Scanned with CamScanner

3. No. of rounds fired from a gun hitting a target.
4. In Radar detection.

ILLUSTRATIVE EXAMPLES

Example 1. (*i*) *Comment on the following statement:*

For a Binomial distribution, mean is 6 and variance is 9.

(*ii*) *A die is tossed thrice. A success is getting 1 or 6 on a toss. Find the mean and variance of the number of success.*

Sol. (*i*)
$$\mu = np = 6 \qquad \text{...(1)}$$
$$\sigma^2 = npq = 9 \qquad \text{...(2)}$$

Dividing (2) by (1), we get
$$q = \frac{9}{6} = 1.5$$

which is impossible as $0 \le q \le 1$

$\therefore$ The above statement is **False.**

(*ii*) Prob. of getting success (1 or 6) on a toss $= \frac{2}{6} = \frac{1}{3} = p$

$$\therefore \qquad q = 1 - \frac{1}{3} = \frac{2}{3}$$

No. of tosses of a die, $n = 3$

(*i*) Mean $= np = 3\left(\frac{1}{3}\right) = 1.$ (*ii*) Variance $= npq = (3)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{3}.$

Example 2. *If 10% of the bolts produced by a machine are defective, determine the probability that out of 10 bolts choosen at random*

(*i*) *1* (*ii*) *None* (*iii*) *at most 2 bolts will be defective.*

Sol. Here, $p(\text{defective}) = \frac{10}{100} = \frac{1}{10}$ (given)

$$\therefore \qquad q(\text{non-defective}) = 1 - \frac{1}{10} = \frac{9}{10}$$

Also, $n = 10$, (n is no. of bolts chosen). (given)

The probability of r defective bolts out of n bolts chosen at random is given by
$$P(r) = {}^nC_r p^r q^{n-r} \qquad \text{...(1)}$$

(*i*) Here $r = 1$,

$$\therefore \qquad P(1) = {}^{10}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{10-1} \qquad | \text{ Using (1)}$$

$$= 10\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)^9 = (.9)^9 = 0.3874 \qquad \text{...(2)}$$

Scanned with CamScanner

(ii) Here $r = 0$

$$\therefore \quad P(0) = {}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10-0} = \left(\frac{9}{10}\right)^{10} = 0.3486 \qquad \text{...(3) | Using (1)}$$

(iii) Prob. that at most 2 bolts will be defective $= P(r \le 2) = P(0) + P(1) + P(2)$ ...(4)

Now,
$$P(2) = {}^{10}C_2 \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{10-2} \qquad \text{| Using (1)}$$

$$= 45\left(\frac{1}{100}\right)(0.43046) = 0.1937$$

$\therefore$ From (4), Required Probability $= P(0) + P(1) + P(2)$

$$= 0.3486 + 0.3874 + 0.1937 = 0.9297.$$

Example 3. *A binomial variable X satisfies the relation $9P(X = 4) = P(X = 2)$ when $n = 6$. Find the value of the parameter p and $P(X = 1)$.*

Sol. We know that

$$P(X = r) = {}^nC_r\, p^r\, q^{n-r} \qquad \text{...(1)}$$

$$\therefore \quad P(X = 4) = {}^6C_4\, p^4 q^2 = 15p^4q^2$$

and
$$P(X = 2) = {}^6C_2 p^2 q^4 = 15\, p^2\, q^4 \qquad \text{| Since } n = 6$$

The given relation is

$$9\,P(X = 4) = P(X = 2) \quad \Rightarrow \quad 9(15p^4 q^2) = 15p^2q^4$$

$$\Rightarrow \quad 9p^2 = q^2 = (1-p)^2 \qquad |\because \ p + q = 1$$

$$\Rightarrow \quad 9p^2 = 1 + p^2 - 2p$$

$$\Rightarrow \quad 8p^2 + 2p - 1 = 0 \quad \Rightarrow \quad (4p-1)(2p+1) = 0$$

$$\therefore \quad \boxed{p = \frac{1}{4}} \qquad |\because \ p \text{ cannot be negative}$$

Now,
$$P(X = 1) = {}^6C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^5 = .3559. \qquad \Big|\because \ q = 1 - p = \frac{3}{4}$$

Example 4. *Fit a binomial distribution to the following frequency data:*

x :	*0*	*1*	*2*	*3*	*4*
f :	*30*	*62*	*46*	*10*	*2*

Sol. The table is as follows:

x	f	fx
0	30	0
1	62	62
2	46	92
3	10	30
4	2	8
	$\Sigma f = 150$	$\Sigma fx = 192$

Mean of observations $= \dfrac{\Sigma fx}{\Sigma f} = \dfrac{192}{150} = 1.28$

$\Rightarrow \quad np = 1.28$

$\Rightarrow \quad 4p = 1.28$ (n is no. of trials)

$\Rightarrow \quad p = 0.32$

$\therefore \quad q = 1 - p = 1 - 0.32 = 0.68$

Also, $\quad N = 150$ $\quad | \because N = \Sigma f$

Hence the binomial distribution is $= N(q+p)^n = 150\,(0.68 + 0.32)^4$.

Example 5. *A student is given a true-false examination with 8 questions. If he corrects at least 7 questions, he passes the examination. Find the probability that he will pass given that he guesses all questions.*

Sol. Here, n = no. of questions asked = 8

$$p = \frac{1}{2}, q = \frac{1}{2} \quad \text{| Since the question can either be true or false}$$

Probability that he will pass

$$= P(r \geq 7) = P(7) + P(8)$$

$$= {}^8C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{8-7} + {}^8C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{8-8} = 8\left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^1 + 1.\left(\frac{1}{2}\right)^8$$

$$= \left(\frac{1}{2}\right)^8 (8+1) = \frac{9}{256} = .03516.$$

Example 6. *During war, 1 ship out of 9 was sunk on an average in making a certain voyage. What was the probability that exactly 3 out of a convoy of 6 ships would arrive safely?*

Sol. p, the probability of a ship arriving safely $= 1 - \dfrac{1}{9} = \dfrac{8}{9}$; $\quad q = \dfrac{1}{9}, n = 6$

The probability that exactly 3 ships arrive safely $= P(r = 3) = {}^6C_3 \left(\dfrac{1}{9}\right)^3 \left(\dfrac{8}{9}\right)^3 = \dfrac{10240}{9^6}$.

Example 7. *A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that dacoit is still alive?*

Sol. Here $\quad n$ = no. of bullets fired = 6, $p = 0.6$, $\quad q = 1 - p = 0.4$

Probability that dacoit is still alive (not killed)

$$= P(r = 0) = {}^nC_0\, p^0\, q^{n-0} = {}^6C_0\,(.6)^0\,(.4)^6 = (.4)^6 = .004096.$$

Example 8. *If the probability of hitting a target is 10% and 10 shots are fired independently. What is the probability that the target will be hit at least once?* (A.K.T.U. 2019)

Sol. Here, $p = \dfrac{10}{100} = \dfrac{1}{10}$, $q = 1 - p = 1 - \dfrac{1}{10} = \dfrac{9}{10}$, $n = 10$

Probability that the target will be hit at least once

$$= P(r \geq 1) = 1 - P(r = 0)$$

$$= 1 - [{}^nC_0\, p^0\, q^n] = 1 - \left[{}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10}\right] = 0.6513.$$

Scanned with CamScanner

Example 9. *Out of 800 families with 4 children each, how many families would be expected to have (i) 2 boys and 2 girls (ii) at least one boy (iii) no girl (iv) atmost two girls? Assume equal probabilities for boys and girls.* (U.P.T.U. 2014)

Sol. Since probabilities for boys and girls are equal,

$$p = \text{probability of having a boy} = \frac{1}{2};\ q = \text{probability of having a girl} = \frac{1}{2}$$

$$n = 4 \qquad\qquad N = 800$$

(*i*) The expected number of families having 2 boys and 2 girls

$$= 800\ {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 800 \times 6 \times \frac{1}{16} = 300.$$

(*ii*) The expected number of families having *at least one boy*

$$= 800\left[{}^4C_1\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right) + {}^4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + {}^4C_3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 + {}^4C_4\left(\frac{1}{2}\right)^4\right]$$

$$= 800 \times \frac{1}{16}\,[4 + 6 + 4 + 1] = 750$$

(*iii*) The expected number of families having no girl *i.e.*, having 4 boys

$$= 800\ .\ {}^4C_4\left(\frac{1}{2}\right)^4 = 50.$$

(*iv*) The expected number of families having *atmost two girls i.e.*, having *at least 2 boy*

$$= 800\left[{}^4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + {}^4C_3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 + {}^4C_4\left(\frac{1}{2}\right)^4\right] = 800 \times \frac{1}{16}\,[6 + 4 + 1] = 550.$$

Example 10. *Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six?*

Sol. p = the chance of getting 5 or 6 with one die $= \frac{2}{6} = \frac{1}{3}$

$$q = 1 - \frac{1}{3} = \frac{2}{3},\ n = 6,\ N = 729$$

Since dice are in sets of 6 and there are 729 sets.

The expected number of times *at least three* dice showing five or six

$$= N\ .\ P(r \geq 3)$$

$$= 729\,[P(3) + P(4) + P(5) + P(6)]$$

$$= 729\left[{}^6C_3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 + {}^6C_4\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4 + {}^6C_5\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5 + {}^6C_6\left(\frac{1}{3}\right)^6\right]$$

$$= \frac{729}{3^6}\,[160 + 60 + 12 + 1] = 233.$$

Example 11. *The probability of a man hitting a target is $\frac{1}{3}$. How many times must he fire so that the probability of his hitting the target at least once is more than 90%?*

Scanned with CamScanner

Sol. $p = \dfrac{1}{3}$

The probability of not hitting the target in n trials is q^n.

Therefore, to find the smallest n for which the probability of hitting at least once is more than 90%, we have

$$1 - q^n > 0.9$$

$$\Rightarrow \qquad 1 - \left(\frac{2}{3}\right)^n > 0.9$$

$$\Rightarrow \qquad \left(\frac{2}{3}\right)^n < 0.1$$

The smallest n for which the above inequality holds true is 6 hence he must fire 6 times.

Example 12. *In a bombing action, there is 50% chance that any bomb will strike the target. Two direct hits are needed to destroy the target completely. How many bombs are required to be dropped to give a 99% chance or better of completely destroying the target?*

Sol. We have, $p = \dfrac{50}{100} = \dfrac{1}{2}$

Since the probability must be greater than 0.99, if n bombs are dropped, we have

$${}^nC_2\left(\frac{1}{2}\right)^n + {}^nC_3\left(\frac{1}{2}\right)^n + {}^nC_4\left(\frac{1}{2}\right)^n + \cdots\cdots + {}^nC_n\left(\frac{1}{2}\right)^n \geq 0.99$$

$$\Rightarrow \qquad \left(\frac{1}{2}\right)^n [{}^nC_2 + {}^nC_3 + {}^nC_4 + \cdots\cdots + {}^nC_n] \geq 0.99$$

$$\frac{2^n - n - 1}{2^n} \geq 0.99$$

$$\Rightarrow \qquad 1 - \frac{1+n}{2^n} \geq 0.99$$

$$\Rightarrow \qquad \frac{1+n}{2^n} \leq 0.01$$

$$\Rightarrow \qquad 2^n\ {}^3\ 100n + 100$$

By trial, $n = 11$ satisfies the inequality.

Hence 11 bombs are required to be dropped.

TEST YOUR KNOWLDGE

1. (*i*) Ten coins are tossed simultaneously. Find the probability of getting at least seven heads.
   (*ii*) A die is thrown five times. If getting an odd number is a success, find the probability of getting at least four successes. *(M.T.U. 2012)*
2. (*a*) The probability of any ship of a company being destroyed on a certain voyage is 0.02. The company owns 6 ships for the voyage. What is the probability of :
   (*i*) losing one ship (*ii*) losing atmost two ships (*iii*) losing none?
   (*b*) Assume that on the average one telephone number out of fifteen called between 2 P.M. and 3 P.M. on week-days is busy. What is the probability that if 6 randomly selected telephone numbers are called (*i*) not more than 3 (*ii*) at least 3 of them will be busy?

Scanned with CamScanner

3. (i) The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of six workers chosen at random, four or more will suffer from the disease?

(ii) The probability that a man aged 60 will live to be 70 is 0.65. What is the probability that out of 10 men, now 60, at least 7 will live to be 70?

4. (i) If the mean of a binomial distribution is 3 and the variance is $\frac{3}{2}$, find the probability of obtaining at least 4 successes.

(ii) In a binomial distribution, for $n = 5$ if $P(x = 1) = 0.4096$ and $P(x = 2) = 0.2048$, then find the value of p.

(iii) The sum and product of the mean and variance of a binomial distribution are $\frac{25}{3}$ and $\frac{50}{3}$ respectively. Find the distribution.

(iv) If the probability of a defective bolt is 0.1, find (a) The mean (b) The standard deviation for the distribution in a total of 400 bolts.

(v) If the moment generating function of a random variable X is $\left(\frac{1}{3}+\frac{2}{3}e^t\right)^5$, find $P(X = 2)$.

5. (a) The probability that a bomb dropped from a plane will strike the target is $\frac{1}{5}$. If six bombs are dropped, find the probability that (i) exactly two will strike the target, (ii) at least two will strike the target.

(b) Four persons in a group of 20 are graduates. If 4 persons are selected at random from 20, find the probability that

(i) all are graduates (ii) at least one is a graduate.

6. A bag contains 5 white, 7 red and 8 black balls. If four balls are drawn, one by one, with replacement, what is the probability that

(i) none is white (ii) all are white
(iii) at least one is white (iv) only 2 are white?

7. (i) In a hurdle race, a player has to cross 10 hurdles. The probability that he will clear each hurdle is $\frac{5}{6}$. What is the probability that he will knock down fewer than 2 hurdles ?

(ii) If on an average one ship in every ten is wrecked, find the probability that out of 5 ships expected to arrive, 4 at least will arrive safely.

8. The prob. that a bulb produced by a factory will fuse after a use of 150 days is 0.05. Find the prob. that out of 5 such bulbs [*M.T.U. (B. Pharma) 2011*]

(i) None (ii) Atmost one
(iii) More than one (iv) At least one will fuse after 150 days of use.

9. Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. What is the probability that

(i) all the five cards are spades (ii) only three are spades? (iii) none is spade?

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

[**Hint.** Manish will take either 6 steps forward and 5 backward or 5 steps forward and 6 backward].

11. (a) In 800 families with 5 children each, how many families would be expected to have (i) 3 boys and 2 girls, (ii) 2 boys and 3 girls, (iii) no girl (iv) at the most two girls. (Assume probabilities for boys and girls to be equal.) (*A.K.T.U. 2017*)

Scanned with CamScanner

(*b*) Out of 800 families with 5 children each, how many would you expect to have (*i*) 3 boys (*ii*) 5 girls (*iii*) either 2 or 3 boys? Assume equal probabilities for boys and girls.

(*M.T.U.* 2012)

(*c*) Out of 320 families with 5 children each, what percentage would be expected to have (*i*) 2 boys and 3 girls (*ii*) at least one boy? Assume equal probability for boys and girls.

(*G.B.T.U.* 2011)

12. In 100 sets of ten tosses of an unbiased coin, in how many cases do you expect to get

(*i*) 7 heads and 3 tails (*ii*) at least 7 heads?

13. The following data are the number of seeds germinating out of 10 on damp filter for 80 sets of seeds. Fit a binomial distribution to this data:

x:	0	1	2	3	4	5	6	7	8	9	10	Total
f:	6	20	28	12	8	6	0	0	0	0	0	80

[**Hint.** Here $n = 10$, $N = 80$, Mean $= \dfrac{\Sigma fx}{\Sigma f} = 2.175$ $\therefore$ $np = 2.175$ etc.]

14. Fit a binomial distribution for the following data and compare the theoretical frequencies with the actual ones :

x:	0	1	2	3	4	5
f:	2	14	20	34	22	8

15. Fit a binomial distribution to the data given in the following table:

(*i*) x:	0	1	2	3	4	
f:	24	41	28	5	2	(*M.T.U.* 2012)
(*ii*) x:	0	1	3	4		
f:	28	62	10	4		(*U.K.T.U.* 2011)

16. (*i*) Assuming half the population of a town consumes chocolates so that the chance of an individual being consumer is $\dfrac{1}{2}$ and that 100 investigators each take 10 individuals to see whether they are consumers, how many investigators would you expect to report that 3 people or less were consumers?

(*ii*) Assuming that 20% of the population of a city are literate, so that the chance of an individual being literate is $\dfrac{1}{5}$ and assuming that 100 investigators each take 10 individuals to see whether they are literate, how many investigators would you expect to report 3 or less were literate?

17. Following results were obtained when 100 batches of seeds were allowed to germinate on damp filter paper in a laboratory : $\beta_1 = \dfrac{1}{15}, \beta_2 = \dfrac{89}{30}$. Determine the Binomial distribution. Calculate the expected frequency for $x = 8$ assuming $p > q$.

18. A coffee connoisseur claims that he can distinguish between a cup of instant coffee and a cup of percolator coffee 75% of the time. It is agreed that his claim will be accepted if he correctly identifies at least 5 of the 6 cups. Find his chances of having the claim (*i*) accepted (*ii*) rejected when he does have the ability he claims.

19. A multiple-choice test consists of 8 questions with 3 answers to each question of which only one is correct. A student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4 and the third answer if he gets 5 or 6. To get a

Scanned with CamScanner

distinction, the student must secure at least 75% correct answers. If there is no negative marking, what is the probability that the student secures a distinction?

20. An irregular six-faced die is thrown and the expectation that in 10 throws, it will give five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws each, would you expect it to give no even number?

Answers

1. (*i*) $\frac{11}{64}$ (*ii*) $\frac{3}{16}$
2. (*a*) (*i*) 0.1085, (*ii*) 0.9997, (*iii*) 0.8858 (*b*) (*i*) 0.9997 (*ii*) 0.005
3. (*i*) $\frac{53}{3125}$ (*ii*) 0.5137
4. (*i*) $\frac{11}{32}$, (*ii*) $\frac{1}{5}$, (*iii*) $\left(\frac{2}{3}+\frac{1}{3}\right)^{15}$

   (*iv*) (*a*) 40 (*b*) 6 (*v*) 0.1646
5. (*a*) (*i*) 0.246 (*ii*) 0.345 (*b*) (*i*) 0.0016 (*ii*) 0.5904
6. (*i*) $\frac{81}{256}$, (*ii*) $\frac{1}{256}$, (*iii*) $\frac{175}{256}$, (*iv*) $\frac{27}{128}$ **7.** (*i*) $\frac{5}{2}\left(\frac{5}{6}\right)^9$, (*ii*) 0.91854
8. (*i*) $\left(\frac{19}{20}\right)^5$ (*ii*) $\frac{6}{5}\left(\frac{19}{20}\right)^4$ (*iii*) $1-\frac{6}{5}\left(\frac{19}{20}\right)^4$ (*iv*) $1-\left(\frac{19}{20}\right)^5$
9. (*i*) $\left(\frac{1}{4}\right)^5$ (*ii*) $90\left(\frac{1}{4}\right)^5$ (*iii*) $\left(\frac{3}{4}\right)^5$ **10.** 0.36787
11. (*a*) (*i*) 250 (*ii*) 250 (*iii*) 25 (*iv*) 400

    (*b*) (*i*) 250 (*ii*) 25 (*iii*) 500

    (*c*) (*i*) 31.25% (*ii*) 96.875%
12. (*i*) 12 nearly (*ii*) 17 nearly **13.** $80\,(0.7825 + 0.2175)^{10}$
14. $100\,(0.432 + 0.568)^5$ **15.** (*i*) $100\,(0.7 + 0.3)^4$, (*ii*) $104\,(0.7404 + 0.2596)^4$
16. (*i*) 17 (*ii*) 88 **17.** $100\left(\frac{1}{4}+\frac{3}{4}\right)^{20}$, 0.075168752
18. (*i*) 0.534 (*ii*) 0.466
19. 0.0197 **20.** 1.

POISSON DISTRIBUTION

4.42 POISSON DISTRIBUTION AS A LIMITING CASE OF BINOMIAL DISTRIBUTION

[*G.B.T.U. 2013 ; M.T.U. 2013, 2014*]

Poisson distribution was discovered by S.D. Poisson in the year 1837.

If the parameters n and p of a binomial distribution are known, we can find the distribution. But in situations where n is very large and p is very small, application of binomial distribution is

Scanned with CamScanner

very labourious. However, if we assume that as $n \to \infty$ and $p \to 0$ such that np always remains finite, say λ, we get the Poisson approximation to the binomial distribution.

Now, for a binomial distribution

$$P(X = r) = {}^nC_r q^{n-r} p^r$$

$$= \frac{n(n-1)(n-2)\dots(n-r+1)}{r!} \times (1-p)^{n-r} \times p^r$$

$$= \frac{n(n-1)(n-2)\dots(n-r+1)}{r!} \times \left(1-\frac{\lambda}{n}\right)^{n-r} \times \left(\frac{\lambda}{n}\right)^r \qquad \Big|\text{ Since } np = \lambda \quad \therefore \quad p = \frac{\lambda}{n}$$

$$= \frac{\lambda^r}{r!} \times \frac{n(n-1)(n-2)\dots(n-r+1)}{n^r} \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\dots\dots\left(\frac{n-r+1}{n}\right) \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\dots\left(1-\frac{r-1}{n}\right) \times \frac{\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^r}$$

As $n \to \infty$, each of the $(r-1)$ factors

$$\left(1-\frac{1}{n}\right), \left(1-\frac{2}{n}\right), \dots\dots, \left(1-\frac{r-1}{n}\right) \qquad \text{tends to 1. Also } \left(1-\frac{\lambda}{n}\right)^r \text{ tends to 1.}$$

Since $\underset{x \to \pm\infty}{\text{Lt}} \left(1+\frac{1}{x}\right)^x = e$, the Naperian base. $\therefore \quad \left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} \longrightarrow e^{-\lambda}$ as $n \to \infty$

Hence in the limiting case when $n \to \infty$, we have

$$\boxed{P(X = r) = \frac{e^{-\lambda}.\lambda^r}{r!} \quad (r = 0, 1, 2, 3, \dots)} \qquad \dots(1)$$

where λ is a finite number $= np$.

(1) represents a probability distribution which is called the *Poisson probability distribution*.

Note 1. λ is called the parameter of the distribution.

Note 2. The sum of the probabilities $P(r)$ for $r = 0, 1, 2, 3, \dots\dots$ is 1, since

$$P(0) + P(1) + P(2) + P(3) + \dots\dots = e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \frac{\lambda^3 e^{-\lambda}}{3!} + \dots\dots$$

$$= e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots\dots\right) = e^{-\lambda}.e^{\lambda} = 1.$$

Scanned with CamScanner

4.43 RECURRENCE FORMULA FOR THE POISSON DISTRIBUTION

For Poisson distribution, $P(r) = \dfrac{e^{-\lambda}\lambda^r}{r!}$ and $P(r+1) = \dfrac{e^{-\lambda}\lambda^{r+1}}{(r+1)!}$

$$\therefore \quad \frac{P(r+1)}{P(r)} = \frac{\lambda r!}{(r+1)!} = \frac{\lambda}{r+1}$$

or
$$P(r+1) = \frac{\lambda}{r+1} P(r), \; r = 0, 1, 2, 3, \ldots\ldots$$

This is called the *recurrence* or *recursion formula* for the Poisson distribution.

4.44 MEAN AND VARIANCE OF THE POISSON DISTRIBUTION

(A.K.T.U. 2018)

For the Poisson distribution, $P(r) = \dfrac{e^{-\lambda}\lambda^r}{r!}$

$$\text{Mean } \mu = \sum_{r=0}^{\infty} rP(r) = \sum_{r=0}^{\infty} r \cdot \frac{e^{-\lambda}\lambda^r}{r!}$$

$$= e^{-\lambda}\sum_{r=1}^{\infty}\frac{\lambda^r}{(r-1)!} = e^{-\lambda}\left(\lambda + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \ldots\ldots\right)$$

$$= \lambda e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots\ldots\right) = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$$

Thus, the mean of the Poisson distribution is equal to the parameter λ.

$$\text{Variance } \sigma^2 = \sum_{r=0}^{\infty} r^2 P(r) - \mu^2 = \sum_{r=0}^{\infty} r^2 \cdot \frac{\lambda^r e^{-\lambda}}{r!} - \lambda^2 = e^{-\lambda}\sum_{r=1}^{\infty}\frac{r^2\lambda^r}{r!} - \lambda^2$$

$$= e^{-\lambda}\left[\frac{1^2 \cdot \lambda}{1!} + \frac{2^2 \cdot \lambda^2}{2!} + \frac{3^2 \cdot \lambda^3}{3!} + \frac{4^2 \cdot \lambda^4}{4!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{2\lambda^2}{1!} + \frac{3\lambda^2}{2!} + \frac{4\lambda^3}{3!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{(1+1)\lambda}{1!} + \frac{(1+2)\lambda^2}{2!} + \frac{(1+3)\lambda^3}{3!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\ldots\right) + \left(\frac{\lambda}{1!} + \frac{2\lambda^2}{2!} + \frac{3\lambda^3}{3!} + \ldots\ldots\right)\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[e^{\lambda} + \lambda\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots\ldots\right)\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\,[e^{\lambda} + \lambda e^{\lambda}] - \lambda^2 = \lambda e^{-\lambda} \cdot e^{\lambda}(1+\lambda) - \lambda^2 = \lambda(1+\lambda) - \lambda^2 = \lambda.$$

Hence, the variance of the Poisson distribution is also λ.

Thus, the mean and the variance of the Poisson distribution are each equal to the parameter λ.

Scanned with CamScanner

Note 1. The mean and the variance of the Poisson distribution can also be derived from those of the binomial distribution in the limiting case when $n \to \infty, p \to 0$ and $np = \lambda$.

Mean of Binomial distribution is np.

$\therefore$ Mean of Poisson distribution $= \underset{n \to \infty}{\text{Lt}}\ np = \underset{n \to \infty}{\text{Lt}}\ \lambda = \lambda$

Variance of Binomial distribution is $npq = np(1-p)$

$\therefore$ Variance of Poisson distribution $= \underset{n \to \infty}{\text{Lt}}\ np(1-p) = \underset{n \to \infty}{\text{Lt}}\ \lambda\left(1 - \frac{\lambda}{n}\right) = \lambda.$

Note 2. For Poisson distribution, $\mu_3 = \lambda$ and $\mu_4 = 3\lambda^2 + \lambda$.

Coefficients of skewness and kurtosis are given by

$$\beta_1 = \frac{1}{\lambda} \text{ and } \gamma_1 = \frac{1}{\sqrt{\lambda}}\text{. Also, } \beta_2 = 3 + \frac{1}{\lambda} \text{ and } \gamma_2 = \frac{1}{\lambda}$$

Hence Poisson distribution is always a skewed distribution.

Remark. While fitting the Poisson distribution to a given data, we round the figures to the nearest integer but it should be kept in mind that the total of the observed and the expected frequencies should be same.

4.45 MODE OF POISSON DISTRIBUTION

Let $$P(x = r) = e^{-\lambda}\frac{\lambda^r}{r!},\ r = 0, 1, 2, \ldots\ldots, \infty$$

The value of r which has a greater probability than any other value is the mode of the Poisson distribution. Thus r is mode if

$$P(X = r) \geq P(X = r + 1) \text{ and } P(X = r) \geq P(X = r - 1)$$

$$\Rightarrow \quad \frac{e^{-\lambda}.\lambda^r}{r!} \geq \frac{e^{-\lambda}.\lambda^{r+1}}{(r+1)!} \quad \text{and} \quad \frac{e^{-\lambda}.\lambda^r}{r!} \geq \frac{e^{-\lambda}.\lambda^{r-1}}{(r-1)!}$$

$$\Rightarrow \quad 1 \geq \frac{\lambda}{r+1} \quad \text{and} \quad \frac{\lambda}{r} \geq 1$$

$$\Rightarrow \quad r \geq \lambda - 1 \quad \text{and} \quad r \leq \lambda \quad \textit{i.e.,} \quad \lambda - 1 \leq r \leq \lambda$$

Case I. If λ is a positive integer, there are two modes $\lambda - 1$ and λ.

Case II. If λ is not a positive integer, there is one mode and is the integral value between $\lambda - 1$ and λ.

4.46 APPLICATIONS OF POISSON DISTRIBUTION

This distribution is applied to problems concerning :

(*i*) Arrival pattern of defective vehicles in a workshop.

(*ii*) Patients in a hospitals.

(*iii*) Telephone calls.

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

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

(*vi*) Emission of radioactive (α) particles.

Scanned with CamScanner

ILLUSTRATIVE EXAMPLES

Example 1. *If the variance of the Poisson distribution is 2, find the probabilities for $r = 1, 2, 3, 4$ from the recurrence relation of the Poisson distribution. Also, find $P(r \geq 4)$.*

(M.T.U. 2013)

Sol. λ, the parameter of Poisson distribution = Variance = 2

Recurrence relation for the Poisson distribution is

$$P(r+1) = \frac{\lambda}{r+1} P(r) = \frac{2}{r+1} P(r) \quad \text{...(1)}$$

Now $$P(r) = \frac{e^{-\lambda} \lambda^r}{r!} \Rightarrow P(0) = \frac{e^{-2}(2)^0}{0!} = e^{-2} = 0.1353$$

Putting $r = 0, 1, 2, 3$ in (1), we get

$$P(1) = 2P(0) = 2 \times 0.1353 = 0.2706 ; \quad P(2) = \frac{2}{2} P(1) = 0.2706$$

$$P(3) = \frac{2}{3} P(2) = \frac{2}{3} \times 0.2706 = 0.1804 ; \quad P(4) = \frac{2}{4} P(3) = \frac{1}{2} \times 0.1804 = 0.0902.$$

Now, $P(r \geq 4) = 1 - [P(0) + P(1) + P(2) + P(3)]$

$= 1 - [0.1353 + 0.2706 + 0.2706 + 0.1804] = 0.1431.$

Example 2. *Using Poisson distribution, find the probability that the ace of spades will be drawn from a pack of well-shuffled cards at least once in 104 consecutive trials.* (U.P.T.U. 2015)

Sol. $$p = \frac{1}{52}, n = 104$$

$\therefore$ $$\lambda = np = 104 \times \frac{1}{52} = 2$$

Prob. (at least once) = $P(r \geq 1) = 1 - P(0)$

$$= 1 - \frac{e^{-\lambda} . \lambda^0}{0!} = 1 - e^{-2} = 1 - 0.135335 \approx 0.8647.$$

Example 3. (*i*) *Fit a Poisson distribution to the following data and calculate theoretical frequencies.*

Deaths:	*0*	*1*	*2*	*3*	*4*
Frequencies:	*122*	*60*	*15*	*2*	*1*

[U.P.T.U. 2014 ; U.K.T.U. 2010]

(*ii*) *The frequency of accidents per shift in a factory is shown in the following table:*

Accident per shift	*Frequency*
0	*192*
1	*100*
2	*24*
3	*3*
4	*1*
Total	**320**

Calculate the mean number of accidents per shift. Fit a Poisson distribution and calculate theoretical frequencies.

Sol. (*i*) Mean of given distribution $= \dfrac{\Sigma fx}{\Sigma f}$

$$\Rightarrow \qquad \lambda = \frac{60 + 30 + 6 + 4}{200} = 0.5$$

Required Poisson distribution $= N \cdot \dfrac{e^{-\lambda} \cdot \lambda^r}{r!} = 200 \cdot \dfrac{e^{-.5}\,(.5)^r}{r!} = (121.306)\dfrac{(0.5)^r}{r!}$

r	$N.\ P(r)$	*Theoretical frequency*
0	$121.306 \times \dfrac{(0.5)^0}{0!} = 121.306$	121
1	$121.306 \times \dfrac{(0.5)^1}{1!} = 60.653$	61
2	$121.306 \times \dfrac{(0.5)^2}{2!} = 15.163$	15
3	$121.306 \times \dfrac{(0.5)^3}{3!} = 2.527$	3
4	$121.306 \times \dfrac{(0.5)^4}{4!} = 0.3159$	0
		Total = 200

(*ii*) Mean number of accidents per shift

$$\lambda = \frac{\Sigma fx}{\Sigma f} = \frac{100 + 48 + 9 + 4}{320} = 0.5031$$

$\therefore$ Required Poisson distribution

$$= N \cdot \frac{e^{-\lambda} \cdot \lambda^r}{r!} = 320 \cdot \frac{e^{-.5031}\,(.5031)^r}{r!} = \frac{(193.48)(.5031)^r}{r!}$$

r	$N.\ P(r)$	*Theoretical frequency*
0	193.48	194
1	97.34	97
2	24.38	24
3	4.10	4
4	0.51	1
		Total = 320

Scanned with CamScanner

Example 4. (*i*) *Suppose that a book of 600 pages contains 40 printing mistakes. Assume that these errors are randomly distributed throughout the book and r, the number of errors per page has a Poisson distribution. What is the probability that 10 pages selected at random will be free from errors?*

(*ii*) *Wireless sets are manufactured with 25 solders joints each, on the average 1 joint in 500 is defective. How many sets can be expected to be free from defective joints in a consignment of 10000 sets?*

Sol. (*i*)

$$p = \frac{40}{600} = \frac{1}{15}, \quad n = 10$$

$$\therefore \quad \lambda = np = 10\left(\frac{1}{15}\right) = \frac{2}{3}$$

$$P(r) = \frac{e^{-\lambda}\,\lambda^r}{r!} = \frac{e^{-2/3}\,(2/3)^r}{r!}$$

$$\therefore \quad P(0) = \frac{e^{-2/3}\,(2/3)^0}{0!} = e^{-2/3} = 0.51.$$

(*ii*)

$$p = \frac{1}{500}, \; n = 25$$

$$\therefore \quad \lambda = np = 25 \times \frac{1}{500} = \frac{1}{20} = 0.05$$

No. of sets in a consignment, N= 10000

Probability of having no defective joint $= P(r = 0) = \dfrac{e^{-.05}\,(0.05)^0}{0!} = 0.9512.$

$\therefore$ The expected no. of sets free from defective joints $= 0.9512 \times 10000 = 9512.$

Example 5. *A manufacturer knows that the condensors he makes contain on an average 1% of defectives. He packes them in boxes of 100. What is the probability that a box picked at random will contain 4 or more faulty condensors?*

Sol.

$$p = 0.01, \; n = 100$$

$$\therefore \quad \lambda = np = 1$$

$$P(r) = \frac{e^{-\lambda}\,\lambda^r}{r!} = \frac{e^{-1}}{r!}$$

P(4 or more faulty condensors) = P(4) + P(5) + ... + P(100)

$$= 1 - [P(0) + P(1) + P(2) + P(3)]$$

$$= 1 - \left[\frac{e^{-1}}{0!} + \frac{e^{-1}}{1!} + \frac{e^{-1}}{2!} + \frac{e^{-1}}{3!}\right]$$

$$= 1 - e^{-1}\left[1 + 1 + \frac{1}{2} + \frac{1}{6}\right] = 1 - \frac{8}{3e} = 1 - 0.981 = 0.019.$$

Example 6. (*i*) *If the probabilities of a bad reaction from a certain injection is 0.0002, determine the chance that out of 1000 individuals more than two will get a bad reaction.*

(*ii*) *The probability that a man aged 50 years will die within a year is 0.01125. What is the probability that of 12 such men, at least 11 will reach their 51st birthday?*

(*Given:* $e^{-.135} = 0.87366$)

Scanned with CamScanner

Sol. (*i*) Here, $p = 0.0002, \quad n = 1000$

$\therefore$ $\lambda = np = 1000 \times 0.0002 = 0.2.$

Since the prob. of bad reaction is very small and no. of trials is very high, we use Poisson distribution here.

The prob. that out of 100 individuals, more than 2 will get a bad reaction is

$$= P(r > 2) = 1 - P(r \leq 2) = 1 - [P(0) + P(1) + P(2)] \quad ...(1)$$

Now, $$P(0) = \frac{e^{-0.2}(0.2)^0}{0} = 0.8187 \quad \text{(Here } r = 0)$$

$$P(1) = \frac{e^{-0.2}(0.2)^1}{1} = 0.1637 \quad \text{(Here } r = 1)$$

and $$P(2) = \frac{e^{-0.2}(0.2)^2}{2} = 0.0164. \quad \text{(Here } r = 2)$$

$\therefore$ From (1), Reqd. probability $= 1 - [0.8187 + 0.1637 + 0.0164] = 0.0012.$

(*ii*) $p = 0.01125, n = 12$

$\therefore$ $\lambda = np = 12 \times 0.01125 = 0.135$

P (at least 11 survive) = P(atmost 1 dies)

$$= P(0) + P(1) = \frac{e^{-\lambda}.\lambda^0}{0!} + \frac{e^{-\lambda}.\lambda^1}{1!}$$

$$= e^{-.135}(1 + 0.135) = 1.135 \times 0.87366 = 0.9916.$$

Example 7. *A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused ($e^{-1.5} = 0.2231$).*

Sol. Since the number of demands for a car is distributed as a Poisson distribution with mean $\lambda = 1.5$.

$\therefore$ Proportion of days on which neither car is used

= Probability of there being no demand for the car

$$= \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-1.5} = 0.2231$$

Proportion of days on which some demand is refused

= probability for the number of demands to be more than two

$$= 1 - P(x \leq 2) = 1 - \left(e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!}\right)$$

$$= 1 - e^{-1.5}\left(1 + 1.5 + \frac{(1.5)^2}{2}\right) = 0.1912625.$$

Example 8. *Suppose the number of telephone calls on an operator received from 9 : 00 to 9 : 05 follow a Poisson distribution with a mean 3. Find the probability that*

(*i*) *The operator will receive no calls in that time interval tomorrow.*

(*ii*) *In the next three days, the operator will receive a total of 1 call in that time interval.*

(Given: $e^{-3} = 0.04978$)

Scanned with CamScanner

Sol. Here, $\lambda = 3$

(i) $$P(0) = \frac{e^{-\lambda}.\lambda^0}{0!} = e^{-3} = 0.04978$$

(ii) Required probability $= P(0)\,P(0)\,P(1) + P(0)\,P(1)\,P(0) + P(1)\,P(0)\,P(0)$

$$= 3\left\{\frac{e^{-\lambda}.\lambda^0}{0!}\right\}^2 \frac{e^{-\lambda}.\lambda^1}{1!} = 9(e^{-3})^3 = 0.00111.$$

Example 9. *The no. of arrivals of customers during any day follows Poisson distribution with a mean of 5. What is the probability that the total no. of customers on two days selected at random is less than 2? (Given: $e^{-10} = 4.54 \times 10^{-5}$)*

Sol. $\lambda = 5$

Arrival of Customers

I day	*II day*	*Total*
0	0	0
0	1	1
1	0	1

Required probability $= P(0)\,P(0) + P(0)\,P(1) + P(1)\,P(0)$

$$= \frac{e^{-5}.5^0}{0!}.\frac{e^{-5}.5^0}{0!} + \frac{e^{-5}.5^0}{0!}.\frac{e^{-5}.5^1}{1!} + \frac{e^{-5}.5^1}{1!}.\frac{e^{-5}.5^0}{0!}$$

$$= e^{-10} + 2\,.\,5\,.\,e^{-10} = 11\,e^{-10} = 11 \times 4.54 \times 10^{-5}$$

$$= 4.994 \times 10^{-4}.$$

Example 10. *An insurance company finds that 0.005% of the population dies from a certain kind of accident each year. What is the probability that the company must pay off no more than 3 of 10,000 insured risks against such incident in a given year?*

Sol. $$p = \frac{0.005}{100} = 0.00005, \quad n = 10000$$

$\therefore$ $$\lambda = np = 10000 \times 0.00005 = 0.5$$

Required Probability $= 1 - P(r \le 3) = 1 - [P(0) + P(1) + P(2) + P(3)]$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{0!} + \frac{e^{-0.5}(0.5)^1}{1!} + \frac{e^{-0.5}(0.5)^2}{2!} + \frac{e^{-0.5}(0.5)^3}{3!}\right]$$

$$= 1 - e^{-5}\,[1 + 0.5 + 0.125 + 0.021] = 0.0016.$$

Example 11. *In a certain factory turning out razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Calculate the approximate number of packets containing no defective, one defective and two defective blades in a consignment of 10,000 packets. (Given: $e^{-0.02} = 0.9802$)*

Sol. $p(\text{defective}) = 0.002$

$n = 10$ (no. of blades in a packet)

$\therefore$ $\lambda = np = 10 \times 0.002 = 0.02$

No. of packets in the consignment, N = 10,000.

Scanned with CamScanner

(*i*) Probability of having no defective = $P(0) = \dfrac{e^{-0.02}\,(0.02)^0}{0\,!} = 0.9802$ | Here $r = 0$

Approximate no. of packets having zero defective in the consignment = 0.9802×10000
$= 9802$

(*ii*) Probability of having one defective = $P(1) = \dfrac{e^{-0.02}\,(0.02)^1}{1\,!} = 0.9802 \times 0.02 = 0.019604$

Approximate no. of packets having one defective in the consignment

$$= 0.019604 \times 10000 \approx 196.$$

(*iii*) Probability of having two defective blades

$$P(2) = \frac{e^{-0.02}\,(0.02)^2}{2\,!} = \frac{(0.980198) \times (0.0004)}{2} = 0.000196.$$

$\therefore$ Approximate no. of packet having two defectives in the consignment

$$= 0.000196 \times 10000 = 1.96 \approx 2.$$

Example 12. (*i*) *Six coins are tossed 6400 times. Using the Poisson distribution, determine the approximate probability of getting six heads x times.*

(*ii*) *A Poisson distribution has a double mode at x = 3 and x = 4. What is the probability that x will have one or the other of these two values?*

Sol. (*i*) Probability of getting one head with one coin $= \dfrac{1}{2}$.

$\therefore$ The probability of getting six heads with six coins $= \left(\dfrac{1}{2}\right)^6 = \dfrac{1}{64}$

$\therefore$ Average number of six heads with six coins in 6400 throws $= np = 6400 \times \dfrac{1}{64} = 100$

$\therefore$ The mean of the Poisson distribution = 100.

Approximate probability of getting six heads x times when the distribution is Poisson

$$= \frac{\lambda^x e^{-\lambda}}{x\,!} = \frac{(100)^x \,.\, e^{-100}}{x\,!}.$$

(*ii*) Since 2 modes are given when λ is an integer, modes are $\lambda - 1$ and λ.

$$\therefore \qquad \lambda - 1 = 3 \quad \Rightarrow \quad \lambda = 4$$

Probability (when $r = 3$) $= \dfrac{e^{-4}\,(4)^3}{3\,!}$

Probability (when $r = 4$) $= \dfrac{e^{-4}\,(4)^4}{4\,!}$

$\therefore$ Required Probability = $P(r = 3 \text{ or } 4) = P(r = 3) + P(r = 4)$

$$= \frac{e^{-4}\,(4)^3}{3\,!} + \frac{e^{-4}\,(4)^4}{4\,!} = \frac{64}{3}\,e^{-4} = 0.39073.$$

Scanned with CamScanner

Example 13. *For a Poisson distribution with mean m, show that*

$$\mu_{r+1} = mr\,\mu_{r-1} + m\frac{d\mu_r}{dm} \text{ where, } \mu_r = \sum_{x=0}^{\infty}(x-m)^r\,\frac{e^{-m}\,.\,m^x}{x!}.$$

Sol.
$$\mu_r = \sum_{x=0}^{\infty}(x-m)^r\,.\,\frac{e^{-m}\,.\,m^x}{x!}$$

$$\frac{d\mu_r}{dm} = \sum_{x=0}^{\infty}\left[\frac{-e^{-m}}{x!}\,.\,m^x\,(x-m)^r + \frac{e^{-m}}{x!}\{xm^{x-1}(x-m)^r - r(x-m)^{r-1}\,.\,m^x\}\right]$$

$$\Rightarrow \quad m\frac{d\mu_r}{dm} = \sum_{x=0}^{\infty}\frac{e^{-m}}{x!}\,m^x\,(x-m)^{r+1} - rm\sum_{x=0}^{\infty}\frac{e^{-m}}{x!}\,m^x\,(x-m)^{r-1} = \mu_{r+1} - mr\,\mu_{r-1}$$

$$\Rightarrow \quad \mu_{r+1} = m\frac{d\mu_r}{dm} + mr\,\mu_{r-1}.$$

Example 14. *Show that in a Poisson distribution with unit mean, mean deviation about mean is* $\left(\frac{2}{e}\right)$ *times the standard deviation.* (G.B.T.U. 2012)

Sol. Here, $\lambda = 1$

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

Mean deviation about mean 1 is

$$= \sum_{x=0}^{\infty}|x-1|\,p(x) = e^{-1}\sum_{x=0}^{\infty}\frac{|x-1|}{x!} = e^{-1}\left[1 + \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \ldots\right] \quad \ldots(1)$$

we have,
$$\frac{n}{(n+1)!} = \frac{\overline{n+1}-1}{(n+1)!} = \frac{1}{n!} - \frac{1}{(n+1)!}$$

$\therefore$ From (1), Mean deviation about mean $= e^{-1}\left[1 + \left(1 - \frac{1}{2!}\right) + \left(\frac{1}{2!} - \frac{1}{3!}\right) + \ldots\right]$

$$= e^{-1}(1+1) = \frac{2}{3} \times 1 = \frac{2}{e} \times \text{S.D.} \qquad |\text{ Since, variance = mean = 1}$$

TEST YOUR KNOWLEDGE

1. If X is a Poisson variate such that P(X = 2) = 9P(X = 4) + 90P(X = 6), find the standard deviation.
2. If a random variable has a Poisson distribution such that P(1) = P(2), find
   (*i*) mean of the distribution (*ii*) P(4).
3. Suppose that X has a Poisson distribution. If $P(X = 2) = \frac{2}{3}P(X = 1)$ find, (*i*) P(X = 0) (*ii*) P(X = 3).

4. A certain screw making machine produces on average 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.

5. (*i*) The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that in a group of 7, five or more will suffer from it?

(*ii*) The experience shows that 4 industrial accidents occur in a plant on an average per month. Calculate the probabilities of less than 3 accidents in a certain month. Use Poisson distribution. (Given : $e^{-4} = 0.01832$). [*M.T.U.* (*MBA*) *2011*]

6. (*i*) Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors ?

(*ii*) Assume that the probability of an individual coalminer being killed in a mine accident during a year is $\frac{1}{2400}$. Use Poisson's distribution to calculate the probability that in a mine employing 200 miners there will be at least one fatal accident in a year.

7. (*i*) A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantee that not more than 10 pins will be defective, what is the approximate prob. that a box will fail to meet the guaranteed quality?

(*ii*) An insurance company insures 4000 people against loss of both eyes in a car accident. Based on previous data it was assumed 10 persons out of 1,00,000 will have such type of injury in car accident. What is probability that more than 2 of the insured will collect on their policy in a given year? (*M.T.U. 2013*)

8. Records show that the probability is 0.00002 that a car will have a flat tyre while driving over a certain bridge. Use Poisson distribution to find the probability that among 20,000 cars driven over the bridge, not more than one will have a flat tyre.

9. Between the hours of 2 and 4 P.M., the average no. of phone calls per minute coming into the switch board of a company is 2.5. Find the probability that during a particular minute, there will be no phone call at all. [Given : $e^{-2} = 0.13534$ and $e^{-0.5} = 0.60650$.]

10. (*i*) Fit a Poisson distribution to the following data given the number of yeast cells per square for 400 squares :

No. of cells per sq. :	0	1	2	3	4	5	6	7	8	9	10
No. of squares :	103	143	98	42	8	4	2	0	0	0	0

It is given that $e^{-1.3225} = 0.2665$.

(*ii*) Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army corps. The distribution of deaths was as follows:

No. of deaths :	0	1	2	3	4	Total
Frequency :	109	65	22	3	1	200

Fit a Poisson distribution to the data and calculate the theoretical frequencies. [*M.T.U.* (*MBA*) *2011*]

(*iii*) The following table gives the no. of days in a 50 day period during which automobile accidents occured in a city.

No. of accidents :	0	1	2	3	4
No. of days :	21	18	7	3	1

Fit a Poisson distribution to the data. (*G.B.T.U. 2011*)

11. The number of accidents in a year involving taxi drivers in a city follows a Poisson distribution with mean equal to 3. Out of 1000 taxi drivers, find approximately the number of drivers with

(*i*) no accidents (*ii*) more than 3 accidents in a year.

Scanned with CamScanner

12. The distribution of the number of road accidents per day in a city is Poisson with mean 4. Find the number of days out of 100 days when there will be

(*i*) no accident (*ii*) at least 2 accidents

(*iii*) atmost 3 accidents (*iv*) between 2 and 5 accidents.

13. It is given that 2% of the electric bulbs manufactured by a company are defective. Using Poisson distribution, find the probability that a sample of 200 bulbs will contain (*i*) no defective bulb (*ii*) two defective bulbs (*iii*) at the most three defective bulbs. (*G.B.T.U. 2011*)

14. A manager accepts the work submitted by his typist only when there is no mistake in the work. The typist has to type on an average 20 letters per day of about 200 words each. Find the chance of her making a mistake if

(*i*) less than 1% of the letters submitted by her are rejected

(*ii*) on 90% days all the letters submitted by her are accepted.

As the probability of making a mistake is small, you may use Poisson distribution. (Take $e = 2.72$)

[**Hint:** (*i*) $e^{-200p} \geq 0.99$ (*ii*) $e^{-4000p} = 0.90$]

15. In a certain factory manufacturing razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Use suitable distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 50,000 packets. (*A.K.T.U. 2018*)

Answers

1. 1 **2.** (*i*) 2 (*ii*) $\dfrac{2}{3e^2}$ **3.** (*i*) $e^{-(4/3)}$ (*ii*) $\dfrac{32}{81}e^{-(4/3)}$ **4.** 0.035

5. (*i*) 0.0008 (*ii*) 0.23816

6. (*i*) 0.4795 (*ii*) 0.08 **7.** (*i*) 0.0136875 (*ii*) 0.007926

8. 0.938 **9.** 0.08208

10. (*i*) Theoretical frequencies are 107, 141, 93, 41, 13, 4, 1, 0, 0, 0, 0.

(*ii*) $108.67 \times \dfrac{(0.61)^r}{r!}$; 109, 66, 20, 4 and 1 (*iii*) $(50)\ \dfrac{e^{-0.9}\,(0.9)^r}{r!}$; 20, 18, 8, 3 and 1.

11. (*i*) 50 (*ii*) 353

12. (*i*) 2 (*ii*) 91 (*iii*) 43 (*iv*) 39

13. (*i*) 0.018315 (*ii*) 0.146525 (*iii*) 0.43347

14. (*i*) $p \leq 0.0000506$ (*ii*) $p = 0.0000263$

15. 49010, 980, 10

4.47 NORMAL DISTRIBUTION

The normal distribution is a continuous distribution. It can be derived from the binomial distribution in the limiting case when n, the number of trials is very large and p, the probability of a success, is close to $\frac{1}{2}$. The general equation of the normal distribution is given by

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

Scanned with CamScanner

4. A certain screw making machine produces on average 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.

5. (*i*) The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that in a group of 7, five or more will suffer from it?

(*ii*) The experience shows that 4 industrial accidents occur in a plant on an average per month. Calculate the probabilities of less than 3 accidents in a certain month. Use Poisson distribution. (Given : $e^{-4} = 0.01832$). [*M.T.U.* (*MBA*) *2011*]

6. (*i*) Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors ?

(*ii*) Assume that the probability of an individual coalminer being killed in a mine accident during a year is $\dfrac{1}{2400}$. Use Poisson's distribution to calculate the probability that in a mine employing 200 miners there will be at least one fatal accident in a year.

7. (*i*) A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantee that not more than 10 pins will be defective, what is the approximate prob. that a box will fail to meet the guaranteed quality?

(*ii*) An insurance company insures 4000 people against loss of both eyes in a car accident. Based on previous data it was assumed 10 persons out of 1,00,000 will have such type of injury in car accident. What is probability that more than 2 of the insured will collect on their policy in a given year? (*M.T.U. 2013*)

8. Records show that the probability is 0.00002 that a car will have a flat tyre while driving over a certain bridge. Use Poisson distribution to find the probability that among 20,000 cars driven over the bridge, not more than one will have a flat tyre.

9. Between the hours of 2 and 4 P.M., the average no. of phone calls per minute coming into the switch board of a company is 2.5. Find the probability that during a particular minute, there will be no phone call at all. [Given : $e^{-2} = 0.13534$ and $e^{-0.5} = 0.60650$.]

10. (*i*) Fit a Poisson distribution to the following data given the number of yeast cells per square for 400 squares :

No. of cells per sq. :	0	1	2	3	4	5	6	7	8	9	10
No. of squares :	103	143	98	42	8	4	2	0	0	0	0

It is given that $e^{-1.3225} = 0.2665$.

(*ii*) Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army corps. The distribution of deaths was as follows:

No. of deaths :	0	1	2	3	4	Total
Frequency :	109	65	22	3	1	200

Fit a Poisson distribution to the data and calculate the theoretical frequencies. [*M.T.U.* (*MBA*) *2011*]

(*iii*) The following table gives the no. of days in a 50 day period during which automobile accidents occured in a city.

No. of accidents :	0	1	2	3	4
No. of days :	21	18	7	3	1

Fit a Poisson distribution to the data. (*G.B.T.U. 2011*)

11. The number of accidents in a year involving taxi drivers in a city follows a Poisson distribution with mean equal to 3. Out of 1000 taxi drivers, find approximately the number of drivers with

(*i*) no accidents (*ii*) more than 3 accidents in a year.

12. The distribution of the number of road accidents per day in a city is Poisson with mean 4. Find the number of days out of 100 days when there will be
(i) no accident (ii) at least 2 accidents
(iii) atmost 3 accidents (iv) between 2 and 5 accidents.

13. It is given that 2% of the electric bulbs manufactured by a company are defective. Using Poisson distribution, find the probability that a sample of 200 bulbs will contain (i) no defective bulb (ii) two defective bulbs (iii) at the most three defective bulbs. (G.B.T.U. 2011)

14. A manager accepts the work submitted by his typist only when there is no mistake in the work. The typist has to type on an average 20 letters per day of about 200 words each. Find the chance of her making a mistake if
(i) less than 1% of the letters submitted by her are rejected
(ii) on 90% days all the letters submitted by her are accepted.
As the probability of making a mistake is small, you may use Poisson distribution. (Take $e = 2.72$)
[**Hint:** (i) $e^{-200p} \geq 0.99$ (ii) $e^{-4000p} = 0.90$]

15. In a certain factory manufacturing razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Use suitable distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 50,000 packets. (A.K.T.U. 2018)

Answers

1. 1 **2.** (i) 2 (ii) $\frac{2}{3e^2}$ **3.** (i) $e^{-(4/3)}$ (ii) $\frac{32}{81}e^{-(4/3)}$ **4.** 0.035
5. (i) 0.0008 (ii) 0.23816`
6. (i) 0.4795 (ii) 0.08 **7.** (i) 0.0136875 (ii) 0.007926
8. 0.938 **9.** 0.08208
10. (i) Theoretical frequencies are 107, 141, 93, 41, 13, 4, 1, 0, 0, 0, 0.
(ii) $108.67 \times \frac{(0.61)^r}{r!}$; 109, 66, 20, 4 and 1 (iii) $(50)\,\frac{e^{-0.9}\,(0.9)^r}{r!}$; 20, 18, 8, 3 and 1.
11. (i) 50 (ii) 353
12. (i) 2 (ii) 91 (iii) 43 (iv) 39
13. (i) 0.018315 (ii) 0.146525 (iii) 0.43347
14. (i) $p \leq 0.0000506$ (ii) $p = 0.0000263$
15. 49010, 980, 10

4.47 NORMAL DISTRIBUTION

The normal distribution is a continuous distribution. It can be derived from the binomial distribution in the limiting case when n, the number of trials is very large and p, the probability of a success, is close to $\frac{1}{2}$. The general equation of the normal distribution is given by

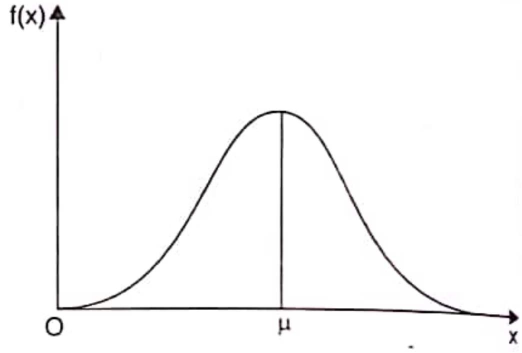
$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

Scanned with CamScanner

where the variable x can assume all values from $-\infty$ to $+\infty$. μ and σ, called the parameters of the distribution, are respectively the mean and the standard deviation of the distribution and $-\infty < \mu < \infty$, $\sigma > 0$. x is called the normal variate and $f(x)$ is called probability density function of the normal distribution.

If a variable x has the normal distribution with mean μ and standard deviation σ, we briefly write $x : N(\mu, \sigma^2)$.

The graph of the normal distribution is called the *normal curve*. It is bell-shaped and symmetrical about the mean μ. The two tails of the curve extend to $+\infty$ and $-\infty$ towards the positive and negative directions of the x-axis respectively and gradually approach the x-axis without ever metting it. The curve is unimodal and the mode of the normal distribution coincides with its mean μ. The line $x = \mu$ divides the area under the normal curve above x-axis into two equal parts. Thus, the median of the distribution also coincides with its mean and mode. The area under the normal curve between any two given ordinates $x = x_1$ and $x = x_2$ represents the probability of values falling into the given interval. The total area under the normal curve above the x-axis is 1.



4.48 BASIC PROPERTIES OF THE NORMAL DISTRIBUTION

The probability density function of the normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

(*i*) $f(x) \geq 0$ (*ii*) $\int_{-\infty}^{\infty} f(x)\, dx = 1$,

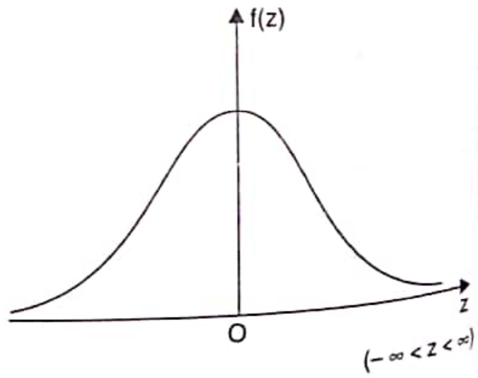
i.e., the total area under the normal curve above the x-axis is 1.

(*iii*) The normal distribution is symmetrical about its mean.

(*iv*) It is a unimodal distribution. The mean, mode and median of this distribution coincide.

4.49 STANDARD FORM OF THE NORMAL DISTRIBUTION

If X is a normal random variable with mean μ and standard deviation σ, then the random variable $Z = \frac{X - \mu}{\sigma}$ has the normal distribution with mean 0 and standard deviation 1. The random variable Z is called the *standardized* (or *standard*) normal random variable.



The probability density function for the normal distribution in standard form is given by

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$

Scanned with CamScanner

It is free from any parameter. This helps us to compute areas under the normal probability curve by making use of standard tables.

Note 1. If $f(z)$ is the probability density function for the normal distribution, then

$$P(z_1 \le Z \le z_2) = \int_{z_1}^{z_2} f(z)\,dz = F(z_2) - F(z_1), \quad \text{where} \quad F(z) = \int_{-\infty}^{z} f(z)\,dz = P(Z \le z)$$

The function $F(z)$ defined above is called the *distribution function* for the normal distribution.

Note 2. The probabilities $P(z_1 \le Z \le z_2)$, $P(z_1 < Z \le z_2)$, $P(z_1 \le Z < z_2)$ and $P(z_1 < Z < z_2)$ are all regarded to be the same.

Note 3. $F(-z_1) = 1 - F(z_1)$.

4.50 NORMAL DISTRIBUTION AS A LIMITING FORM OF BINOMIAL DISTRIBUTION (when $p = q$)

Let $N(q+p)^n$ be the binomial distribution. If $p = q$ then $p = q = \frac{1}{2}$ (since $p + q = 1$) and consequently the binomial distribution is symmetrical. Let n be an even integer say $2k$, k being an integer. Since $n \to \infty$, the frequencies of r and $r + 1$ successes can be written in following forms:

$$f(r) = N.\, {}^{2k}C_r \left(\frac{1}{2}\right)^{2k}$$

$$f(r+1) = N.\, {}^{2k}C_{r+1} \left(\frac{1}{2}\right)^{2k}$$

$$\therefore \quad \frac{f(r+1)}{f(r)} = \frac{{}^{2k}C_{r+1}}{{}^{2k}C_r} = \frac{2k - r}{r+1}$$

The frequency of r successes will be greater than the frequency of $(r + 1)$ successes if

$$f(r) > f(r+1)$$

$$\Rightarrow \quad \frac{f(r+1)}{f(r)} < 1$$

$$\Rightarrow \quad 2k - r < r + 1$$

$$\Rightarrow \quad r > k - \frac{1}{2} \qquad \ldots(1)$$

In a similar way, the frequency of r successes will be greater than the frequencies of $(r-1)$ successes if $r < k + \frac{1}{2}$ ...(2)

In view of (1) and (2), we observe that if $k - \frac{1}{2} < r < k + \frac{1}{2}$ the frequency corresponding to r successes will be the greatest. Clearly, $r = k$ is the value of the success corresponding to which the frequency is maximum. Suppose it is y_0. Then, we have

$$y_0 = N.\, {}^{2k}C_k \left(\frac{1}{2}\right)^{2k} = N.\frac{2k!}{k!\,k!}\left(\frac{1}{2}\right)^{2k}$$

Let y_x be the frequency of $k + x$ successes then, we have

$$y_x = N \cdot {}^{2k}C_{k+x} \left(\frac{1}{2}\right)^{2k} = N \cdot \left(\frac{1}{2}\right)^{2k} \cdot \frac{2k!}{(k+x)!\,(k-x)!}$$

Now,
$$\frac{y_x}{y_0} = \frac{k!\,k!}{(k+x)!\,(k-x)!} = \frac{k(k-1)(k-2)\ldots(k-x+1)}{(k+x)(k+x-1)\ldots(k+1)}$$

$$= \frac{\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\ldots\left\{1-\frac{x-1}{k}\right\}}{\left(1+\frac{1}{k}\right)\left(1+\frac{2}{k}\right)\ldots\left(1+\frac{x}{k}\right)}$$

Taking log on both sides,

$$\log\frac{y_x}{y_0} = \left[\log\left(1-\frac{1}{k}\right) + \log\left(1-\frac{2}{k}\right) + \cdots + \log\left(1-\frac{x-1}{k}\right)\right]$$

$$-\left[\log\left(1+\frac{1}{k}\right) + \log\left(1+\frac{2}{k}\right) + \cdots + \log\left(1+\frac{x}{k}\right)\right] \quad \text{...(3)}$$

Now, writing expression for each term and neglecting higher powers of $\frac{x}{k}$ (very small quantity), we get from (3),

$$\log\frac{y_x}{y_0} = -\frac{1}{k}\{1+2+3+\cdots+(x-1)\} - \frac{1}{k}\{1+2+3+\ldots+(x-1)+x\}$$

$$= -\frac{2}{k}\{1+2+3+\cdots+(x-1)\} - \frac{x}{k}$$

$$= -\frac{2}{k}\,\frac{(x-1)\,x}{2} - \frac{x}{k} = -\frac{x^2}{k}$$

$\therefore$
$$y_x = y_0\, e^{-x^2/k}$$

$\Rightarrow$
$$\boxed{y_x = y_0\, e^{-x^2/2\sigma^2}} \qquad \Big| \because \quad \sigma^2 = npq = \frac{n}{4} = \frac{k}{2}$$

which is **normal distribution.**

4.51 MEAN AND VARIANCE OF NORMAL DISTRIBUTION (A.K.T.U. 2015, 2018)

1. The A.M. of a continuous distribution $f(x)$ is given by

$$\text{A.M.}\ (\bar{x}) = \frac{\int_{-\infty}^{\infty} x\, f(x)\, dx}{\int_{-\infty}^{\infty} f(x)\, dx} \qquad \Big| \text{ By definition}$$

Consider the normal distribution with μ, σ as the parameters then

$$\bar{x} = \int_{-\infty}^{\infty} x.\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad \Bigg| \text{ Since } \int_{-\infty}^{\infty} f(x)\,dx = \text{area under normal curve} = 1$$

Put $\frac{x-\mu}{\sigma} = z$ so that $x = \mu + \sigma z \quad \therefore \quad dx = \sigma dz$

so,

$$\bar{x} = \int_{-\infty}^{\infty} (\mu + \sigma z)\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} (\sigma\, dz)$$

$$= \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2}\, dz$$

$$= \mu + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, d\left(\frac{z^2}{2}\right) \quad \Bigg| \because \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 1$$

$$= \mu + \frac{\sigma}{\sqrt{2\pi}} \left(\frac{e^{-z^2/2}}{-1}\right)_{-\infty}^{\infty}$$

$$\boxed{\bar{x} = \mu}$$

2. By definition,

$$\text{Variance} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx + \bar{x}^2 \int_{-\infty}^{\infty} f(x)\, dx - 2\bar{x} \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx + \bar{x}^2 - 2\bar{x}\,\bar{x} \quad \Bigg| \because \int_{-\infty}^{\infty} f(x)\, dx = 1 \text{ and } \int_{-\infty}^{\infty} x f(x)\, dx = \bar{x}$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx - \bar{x}^2 \qquad ...(1)$$

Now, Let $I = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} dx$

Put $\frac{x-\bar{x}}{\sigma} = z$ so that $x = \bar{x} + \sigma z \quad \therefore \quad dx = \sigma\, dz$

Hence,

$$I = \int_{-\infty}^{\infty} (\bar{x} + \sigma z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2}\, \sigma\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \left[\sigma^2 \int_{-\infty}^{\infty} z^2 e^{-z^2/2}\, dz + \bar{x}^2 \int_{-\infty}^{\infty} e^{-z^2/2}\, dz + 2\sigma\bar{x} \int_{-\infty}^{\infty} z e^{-z^2/2}\, dz\right]$$

$$= \frac{-\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, d(e^{-z^2/2}) + \bar{x}^2 .1 + 2\sigma\bar{x} .0$$

Scanned with CamScanner

$$= -\frac{\sigma^2}{\sqrt{2\pi}}\left(ze^{-z^2/2}\right)_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}\, dz + \bar{x}^2$$

$$= 0 + \sigma^2 . 1 + \bar{x}^2 = \sigma^2 + \bar{x}^2$$

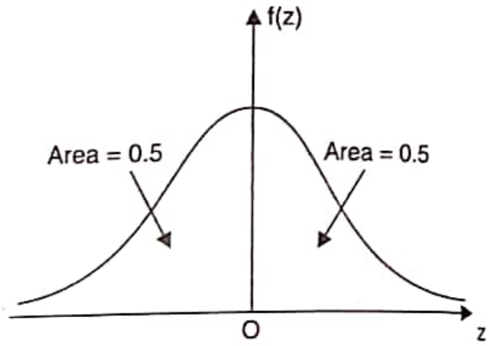
$\therefore$ From (1), $\boxed{\text{Variance} = \sigma^2 + \bar{x}^2 - \bar{x}^2 = \sigma^2}$

$\therefore$ The standard deviation of the normal distribution is σ.

4.52 AREA UNDER THE NORMAL CURVE

By taking $z = \frac{x-\mu}{\sigma}$, standard normal curve is formed. The total area under this curve is 1.

The area under the curve is divided into two equal parts by $z = 0$. The area between the ordinate $z = 0$ and any other ordinate can be noted from the supplied table. It should be noted that mean for the normal distribution is 0.



4.53 APPLICATIONS OF NORMAL DISTRIBUTION

De Moivre made the discovery of this distribution in 1733.

This distribution has an important application in the theory of errors made by chance in experimental measurements. Its more applications are in computation of hit probability of a shot and in statistical inference in almost every branch of science.

ILLUSTRATIVE EXAMPLES

Example 1. *A sample of 100 dry battery cells tested to find the length of life produced the following results:*

$$\bar{x} = 12 \text{ hours}, \sigma = 3 \text{ hours}.$$

Assuming the data to be normally distributed, what percentage of battery cells are expected to have life

(i) more than 15 hours *(ii) less than 6 hours* *(iii) between 10 and 14 hours?*

(A.K.T.U. 2018)

Sol. Here x denotes the length of life of dry battery cells.

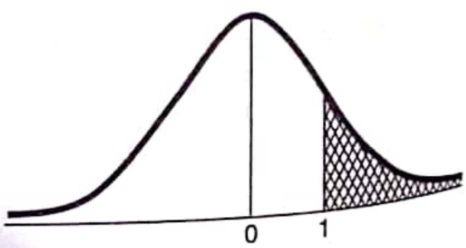
Also, $z = \frac{x-\bar{x}}{\sigma} = \frac{x-12}{3}$.

(*i*) When $x = 15, z = 1$

$\therefore$ $P(x > 15) = P(z > 1)$

$= P(0 < z < \infty) - P(0 < z < 1)$

$= .5 - 0.3413 = 0.1587 = 15.87\%$.



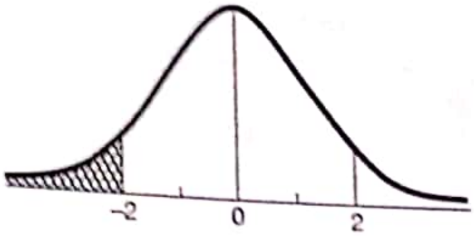
Scanned with CamScanner

(*ii*) When $x = 6$, $z = -2$

$\therefore$ $P(x < 6) = P(z < -2)$

$= P(z > 2) = P(0 < z < \infty) - P(0 < z < 2)$

$= 0.5 - 0.4772 = 0.0228 = 2.28\%$.



(*iii*) When $x = 10$, $z = -\dfrac{2}{3} = -0.67$

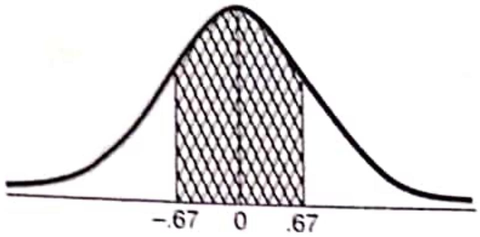
When $x = 14$, $z = \dfrac{2}{3} = 0.67$

$P(10 < x < 14)$

$= P(-0.67 < z < 0.67)$

$= 2P(0 < z < 0.67) = 2 \times 0.2485$

$= 0.4970 = 49.70\%$.



Example 2. *In a sample of 1000 cases, the mean of a certain test is 14 and S.D. is 2.5. Assuming the distribution to be normal, find*

(*i*) *how many students score between 12 and 15?*

(*ii*) *how many score above 18?*

(*iii*) *how many score below 8?*

(*iv*) *how many score 16?*

Sol. (*i*) $$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{12 - 14}{2.5} = -0.8$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{15 - 14}{2.5} = 0.4$$

Area lying between -0.8 and 0.4

= Area between 0 to 0.8 + Area between 0 to 0.4

$= 0.2881 + 0.1554 = 0.4435$

Reqd. no. of students $= 1000 \times 0.4435 = 444$ (app.)

(*ii*) $$z = \frac{18 - 14}{2.5} = 1.6$$

Area right to $1.6 = 0.5 -$ (Area between 0 and 1.6) $= 0.5 - 0.4452 = 0.0548$

Reqd. no. of students $= 1000 \times 0.0548 = 54.8 \approx 55$ (app.)

(*iii*) $$z = \frac{8 - 14}{2.5} = -2.4$$

Area left to $-2.4 = 0.5 -$ (Area between 0 and 2.4) $= 0.5 - 0.4918 = 0.0082$

$\therefore$ Reqd. no. of students $= 1000 \times 0.0082 = 8.2 \approx 8$ (app.).

(*iv*) $$z_1 = \frac{15.5 - 14}{2.5} = 0.6$$

$$z_2 = \frac{16.5 - 14}{2.5} = 1$$

Area between 0.6 and 1 $= 0.3413 - 0.2257 = 0.1156$

$\therefore$ Reqd. no. of students $= 1000 \times 0.1156 = 115.6 \approx 116$ (app.).

Scanned with CamScanner

Example 3. *Assume mean height of soldiers to be 68.22 inches with a variance of 10.8 inches square. How many soldiers in a regiment of 1,000 would you expect to be over 6 feet tall, given that the area under the standard normal curve between z = 0 and z = 0.35 is 0.1368 and between z = 0 and z = 1.15 is 0.3746.* [G.B.T.U. (C.O.) 2011]

Sol. x = 6 feets = 72 inches

$$\therefore \qquad z = \frac{x-\mu}{\sigma} = \frac{72-68.22}{\sqrt{10.8}} = 1.15$$

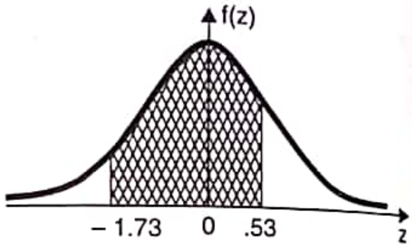
$$P(x > 72) = P(z > 1.15) = 0.5 - P(0 \le z \le 1.15)$$
$$= 0.5 - 0.3746 = 0.1254$$

$\therefore$ Expected no. of soldiers = 1000 × 0.1254 = 125.4 ≈ 125 (app.).

Example 4. *A large number of measurement is normally distributed with a mean 65.5" and S.D. of 6.2". Find the percentage of measurements that fall between 54.8" and 68.8".*

Sol. Mean μ = 65.5 inches, S.D. σ = 6.2 inches

x_1 = 54.8 inches, x_2 = 68.8 inches



$$\therefore \qquad z_1 = \frac{x_1-\mu}{\sigma} = \frac{54.8-65.5}{6.2} = -1.73$$

and
$$z_2 = \frac{x_2-\mu}{\sigma} = \frac{68.8-65.5}{6.2} = 0.53$$

Now, $P(-1.73 \le z \le 0.53) = P(-1.73 \le z \le 0) + P(0 \le z \le 0.53)$
$= P(0 \le z \le 1.73) + P(0 \le z \le 0.53)$
$= 0.4582 + 0.2019 = 0.6601$ | By table

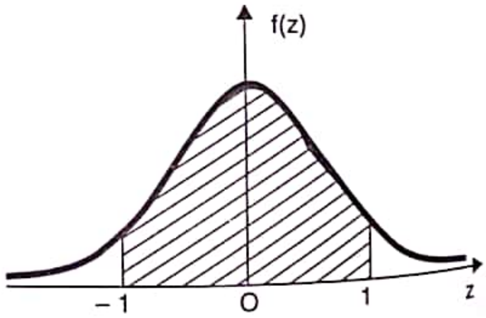
$\therefore$ Reqd. percentage of measurements = 66.01%.

Example 5. *A manufacturer knows from experience that the resistance of resistors he produces is normal with mean μ = 100 ohms and standard deviation σ = 2 ohms. What percentage of resistors will have resistance between 98 ohms and 102 ohms?*

Sol. $\mu = 100\ \Omega, \quad \sigma = 2\ \Omega, \quad x_1 = 98\ \Omega, \quad x_2 = 102\ \Omega$

$$\therefore \qquad z_1 = \frac{x_1-\mu}{\sigma} = \frac{98-100}{2} = -1$$

and
$$z_2 = \frac{x_2-\mu}{\sigma} = \frac{102-100}{2} = 1.$$



Now, $P(98 < x < 102) = P(-1 < z < 1)$
$= P(-1 \le z \le 0) + P(0 \le z \le 1)$
$= P(0 \le z \le 1) + P(0 \le z \le 1)$
$= 0.3413 + 0.3413 = 0.6826.$

$\therefore$ Percentage of resistors having resistance between 98 Ω and 102 Ω = 68.26%.

Example 6. *In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution. It is given that if $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{1}{2}x^2} dx$ then f(0.5) = 0.19 and f(1.4) = 0.42.* (A.K.T.U. 2019, 2018)

Sol. Let μ and σ be the mean and S.D. respectively.

31% of the items are under 45.

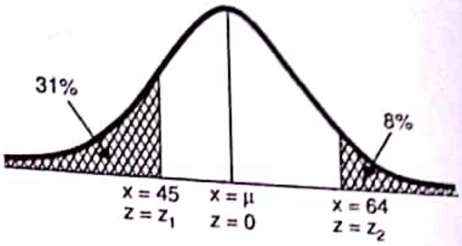
$\Rightarrow$ Area to the left of the ordinate $x = 45$ is 0.31

When $x = 45$, let $z = z_1$

$$P(z_1 < z < 0) = 0.5 - 0.31 = 0.19$$

From the tables, the value of z corresponding to this area is 0.5

$\therefore$ $$z_1 = -0.5 [z_1 < 0]$$



When $x = 64$, let $z = z_2$

$$P(0 < z < z_2) = 0.5 - 0.08 = 0.42$$

From the tables, the value of z corresponding to this area is 1.4.

$\therefore$ $$z_2 = 1.4$$

Since, $$z = \frac{x - \mu}{\sigma}$$

$\therefore$ $$-0.5 = \frac{45 - \mu}{\sigma} \quad \text{and} \quad 1.4 = \frac{64 - \mu}{\sigma}$$

$\Rightarrow$ $$45 - \mu = -0.5\sigma \qquad ...(1)$$

and $$64 - \mu = 1.4\sigma \qquad ...(2)$$

Subtracting $-19 = -1.9\sigma \quad \therefore \quad \sigma = 10$

From (1), $45 - \mu = -0.5 \times 10 = -5 \quad \therefore \quad \mu = 50.$

Example 7. *The life of army shoes is normally distributed with mean 8 months and standard deviation 2 months. If 5000 pairs are insured, how many pairs would be expected to need replacement after 12 months?* $\left[\textit{Given that } P(z \geq 2) = 0.0228 \textit{ and } z = \frac{x - \mu}{\sigma}\right]$.

(A.K.T.U. 2018)

Sol. Mean $(\mu) = 8$, Standard Deviation $(\sigma) = 2$

Number of pairs of shoes = 5000, Total months $(x) = 12$

when $x = 12$, $$z = \frac{x - \mu}{\sigma} = \frac{12 - 8}{2} = 2$$

Area $(z \geq 2) = 0.0228$

Number of pairs whose life is more than 12 months = 5000 × 0.0228 = 114

Pair of shoes needing replacement after 12 months = 5000 – 114 = 4886.

Example 8. *The mean inside diameter of a sample of 200 washers produced by a machine is 0.502 cm and the standard deviation is 0.005 cm. The purpose for which these washers are intended allows a minimum tolerance in the diameter of 0.496 to 0.508 cm, otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine. Assume the diameters are normally distributed.*

Sol. Given: Mean $\mu = 0.502$ cm, S.D. $\sigma = 0.005$ cm, $x_1 = 0.496$ cm, $x_2 = 0.508$ cm.

Now, $$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{0.496 - 0.502}{0.005} = -1.2$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{0.508 - 0.502}{0.005} = 1.2$$

Scanned with CamScanner

Area for non-defective washers

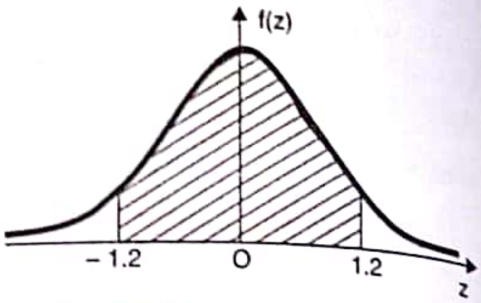
$= P(-1.2 \le z \le 1.2)$

$= P(-1.2 \le z \le 0) + P(0 \le z \le 1.2)$

$= P(0 \le z \le 1.2) + P(0 \le z \le 1.2)$

$= 0.3849 + 0.3849 = 0.7698$

$= 76.98\%.$



$\therefore$ Percentage of defective washers $= 100 - 76.98$

$= 23.02\%.$

Example 9. *Assuming that the diameters of 1000 brass plugs taken consecutively from a machine, form a normal distribution with mean 0.7515 cm and standard deviation 0.002 cm, how many of the plugs are likely to be rejected if the approved diameter is 0.752 ± 0.004 cm.*

Sol. Tolerance limits of the diameter of non-defective plugs are

$0.752 - 0.004 = 0.748$ cm. and $0.752 + 0.004 = 0.756$ cm.

Standard normal variable, $z = \dfrac{x - \mu}{\sigma}$

If $x_1 = 0.748$, $z_1 = \dfrac{0.748 - 0.7515}{0.002} = -1.75$

If $x_2 = 0.756$, $z_2 = \dfrac{0.756 - 0.7515}{0.02} = 2.25$

Area from ($z_1 = -1.75$) to ($z_2 = 2.25$)

$= P(-1.75 \le z \le 2.25) = P(-1.75 \le z \le 0) + P(0 \le z \le 2.25)$

$= P(0 \le z \le 1.75) + P(0 \le z \le 2.25) = 0.4599 + 0.4878 = 0.9477$

Number of plugs which are likely to be rejected $= 1000 \times (1 - 0.9477) = 1000 \times .0523 = 52.3$

Hence approximately 52 plugs are likely to be rejected.

Example 10. *If the heights of 300 students are normally distributed with mean 64.5 inches and standard deviation 3.3 inches, find the height below which 99% of the students lie.*

Sol. Mean $\mu = 64.5$ inches, S.D. $\sigma = 3.3$ inches

Area between 0 and $\dfrac{x - 64.5}{3.3} = 0.99 - 0.5 = 0.49$

From the table, for the area 0.49, $z = 2.327$

The corresponding value of x is given by

$$\frac{x - 64.5}{3.3} = 2.327$$

$\Rightarrow \quad x - 64.5 = 7.68$

$\Rightarrow \quad x = 7.68 + 64.5 = 72.18$ inches.

Hence 99% students are of height less than 6 ft. 0.18 inches.

Example 11. *The income of a group of 10,000 persons was found to be normally distributed with mean ₹ 750 p.m. and standard deviation of ₹ 50. Show that, of this group, about 95% had income exceeding ₹ 668 and only 5% had income exceeding ₹ 832. Also find the lowest income among the richest 100.*

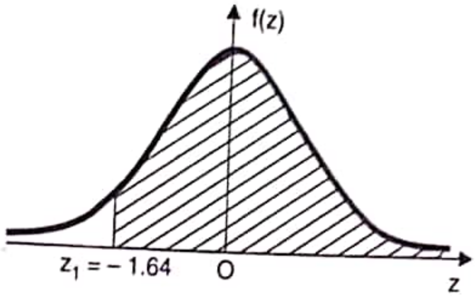
Sol. Given: $\mu = 750, \sigma = 50$

Scanned with CamScanner

Standard normal variable, $z = \dfrac{x-\mu}{\sigma}$

(*i*) If $x_1 = 668$, $z_1 = \dfrac{x_1-\mu}{\sigma} = \dfrac{668-750}{50} = -1.64$

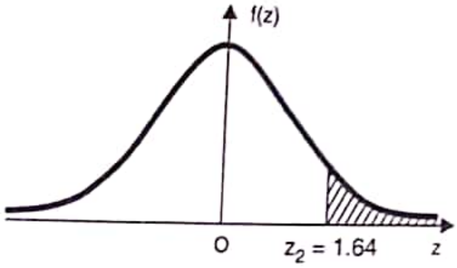
$$\begin{aligned} P(x_1 > 668) &= P(z_1 > -1.64) \\ &= 0.5 + P(-1.64 \le z \le 0) \\ &= 0.5 + P(0 \le z \le 1.64) \\ &= 0.5 + 0.4495 \\ &= 0.9495 \end{aligned}$$



$\therefore$ Required percentage of persons having income exceeding ₹ 668 = 94.95% ≈ 95% (approx.)

(*ii*) If $x_2 = 832$, $z_2 = \dfrac{x_2-\mu}{\sigma} = \dfrac{832-750}{50} = 1.64$

$$\begin{aligned} P(x_2 > 832) &= P(z_2 > 1.64) \\ &= 0.5 - P(0 \le z \le 1.64) \\ &= 0.5 - 0.4495 = 0.0505 \end{aligned}$$



$\therefore$ Required percentage of persons having income exceeding ₹ 832 = 5.05% ≈ 5% (approx.)

(*iii*) Let x be the lowest income among the richest 100 persons *i.e.*, 1% of 10,000.

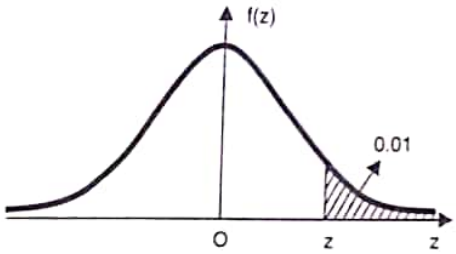
Thus, area between O and $z = 0.49$ (see figure) by Normal distribution table,

$$z = 2.33$$

Thus, $\dfrac{x-\mu}{\sigma} = 2.33$

$\Rightarrow$ $\dfrac{x-750}{50} = 2.33$

$\Rightarrow$ $x = 866.5$



Hence ₹ 866.5 is the minimum income among the richest 100 persons.

Example 12. *255 metal rods were cut roughly 6 inches over size. Finally the lengths of the over size amount, were measured exactly and grouped with 1 inch intervals, there being in all 12 groups* $\frac{1}{2}'' - 1\frac{1}{2}''$, $1\frac{1}{2}'' - 2\frac{1}{2}''$, ..., $11\frac{1}{2}'' - 12\frac{1}{2}''$.

The frequency distribution for the 255 lengths was as follows:

Length (inches) Central value	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Frequency	*2*	*10*	*19*	*25*	*40*	*44*	*41*	*28*	*25*	*15*	*5*	*1*

Fit a normal curve to this data.

Sol. The equation of the normal curve for N observations is

$$y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \text{...(1)}$$

x	f	$u = x - 6$	fu	fu^2
1	2	– 5	– 10	50
2	10	– 4	– 40	160
3	19	– 3	– 57	171
4	25	– 2	– 50	100
5	40	– 1	– 40	40
6	44	0	0	0
7	41	1	41	41
8	28	2	56	112
9	25	3	75	225
10	15	4	60	240
11	5	5	25	125
12	1	6	6	36
Total	255		66	1300

Mean, $\mu = a + \dfrac{\Sigma fu}{\Sigma f} = 6 + \dfrac{66}{255} = 6.259$

Variance, $\sigma^2 = \dfrac{\Sigma fu^2}{\Sigma f} - \left(\dfrac{\Sigma fu}{\Sigma f}\right)^2 = \dfrac{1300}{225} - \left(\dfrac{66}{255}\right)^2 = 5.031$

$\therefore$ $\sigma = 2.243$

Thus, we have N = 255, Mean, $\mu = 6.259''$, S.D. $\sigma = 2.243''$

Hence the fitted curve is

$$y = \frac{255}{2.243\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x - 6.259}{2.243}\right)^2} \qquad | \text{ From (1)}$$

$$= \frac{113.68}{\sqrt{2\pi}} e^{-0.099(x-6.259)^2}.$$

Example 13. *Show that the area under the normal curve is unity.*

Sol. Area under the normal curve is given by

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

Put $\dfrac{x - \mu}{\sigma} = z$ so $dx = \sigma\, dz$

$$\therefore \quad A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} (\sigma dz) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-z^2/2}\, dz$$

Now, $$A \cdot A = A^2 = \left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-x^2/2}\, dx\right)\left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-y^2/2}\, dy\right)$$

$$= \frac{2}{\pi} \int_0^{\infty}\int_0^{\infty} e^{-\left(\frac{x^2+y^2}{2}\right)} dx\, dy \qquad | \text{ where } x \text{ and } y \text{ are dummy variables}$$

Scanned with CamScanner

Put $x = r\cos\theta$, $y = r\sin\theta$ so that $J = r$ changing to polar coordinates,

$$A^2 = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{\infty} e^{-r^2/2}\, r\, dr\, d\theta = \int_0^{\infty} e^{-r^2/2}\, d\left(\frac{r^2}{2}\right) = 1$$

$\therefore$ A = Area under the normal curve = 1

Example 14. *Prove that for normal distribution, the mean deviation from the mean equals to $\frac{4}{5}$ of the standard deviation approximately.*

Sol. Let μ and σ be the mean and standard deviation of the normal distribution. Then by definition,

Mean deviation from the mean

$$= \int_{-\infty}^{\infty} |x-\mu|\, f(x)\, dx$$

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

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \sigma|z|\, e^{-\frac{1}{2}z^2}\, \sigma\, dz \qquad \left|\ \begin{array}{l}\text{when } \dfrac{x-\mu}{\sigma} = z\\ \Rightarrow\ dx = \sigma\, dz\end{array}\right.$$

$$= \sigma\sqrt{\frac{2}{\pi}}\int_0^{\infty} z\, e^{-z^2/2}\, dz$$

$$= \sigma\sqrt{\frac{2}{\pi}}\left[-e^{-z^2/2}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\,\sigma = 0.7979\,\sigma \approx 0.8\sigma \approx \frac{4}{5}\sigma$$

TEST YOUR KNOWLEDGE

1. In a test on 2000 electric bulbs, it was found that the life of a particular make, was normally distributed with an average life of 2040 hours and S.D. of 60 hours, estimate the number of bulbs likely to burn for
   (*i*) more than 2150 hours (*ii*) less than 1950 hours
   (*iii*) more than 1920 hours but less than 2160 hours.
2. An aptitude test for selecting officers in a bank is conducted on 1000 candidates. The average score is 42 and the standard deviation of score is 24. Assuming normal distribution for the scores, find
   (*i*) the number of candidates whose scores exceed 60.
   (*ii*) the number of candidates whose scores lie between 30 and 60.
3. (*i*) In a normal distribution exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution?
   (*ii*) In a normal distribution, 0.0107 of the items lie below 42 and 0.0446 of the items lie above 82. What is the mean and standard deviation of the normal distribution?
4. If Z is a standard normal variable, find the following probabilities:
   (*i*) P(Z < 1.2) (*ii*) P(Z > – 1.2) (*iii*) P(– 1.2 < Z < 1.3).

Scanned with CamScanner

5. An aptitude test was conducted on 900 employees of the Metro Tyres Limited in which the mean score was found to be 50 units and standard deviation was 20. On the basis of this information, you are required to answer the following questions:
(i) What was the number of employees whose mean score was less than 30?
(ii) What was the number of employees whose mean score exceeded 70?
(iii) What was the number of employees whose mean score were between 30 and 70?

$\frac{x-\mu}{\sigma}$	0.25	0.50	0.70	1.00	1.25	1.50
Area	0.0987	0.1915	0.2734	0.3413	0.3944	0.4332

[U.P.T.U. (MBA) 2009]

6. (a) Students of a class were given a mechanical aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percent of students scored?
(i) more than 60 marks? (ii) less than 56 marks? (iii) between 45 and 65 marks?
(b) 2000 students appeared in an examination. Distribution of marks is assumed to be normal with mean $\mu = 30$ and $\sigma = 6.25$. How many students are expected to get marks?
(i) between 20 and 40 (ii) less than 35 and (iii) above 50.
[U.P.T.U. (MBA) 2012]
(c) Suppose the weight W of 600 male students are normally distributed with mean $\mu = 70$ kg and standard deviation $\sigma = 5$ kg. Find number of students with weight
(i) between 69 and 74 kg (ii) more than 76 kg. (G.B.T.U. 2013)

7. (a) In an intelligence test administered to 1000 students, the average score was 42 and standard deviation 24. Find:
(i) the expected number of students scoring more than 50.
(ii) the number of students scoring between 30 and 54.
(iii) the value of score exceeded by top 100 students. [G.B.T.U. (MBA) 2010]
(b) The average monthly sales of 5000 firms are normally distributed. Its mean and standard deviation are ₹ 36000 and ₹ 10000 respectively. Find:
(i) the no. of firms having sales over ₹ 40000.
(ii) the no. of firms having sales between ₹ 30000 and ₹ 40000.
[Given area under normal curve from 0 to z for Z (0.4) = 0.1554 and Z (0.6) = 0.2257]
[G.B.T.U. (MBA) 2010]
(c) The daily wages of 1000 workers are distributed around a mean of ₹ 140 and with a standard deviation of ₹ 10. Estimate the number of workers whose daily wages will be
(i) between ₹ 140 and ₹ 144 (ii) less than ₹ 126
(iii) more than ₹ 160. (G.B.T.U. 2012)

8. (a) Records kept by the goods inwards department of a large factory show that the average no. of lorries arriving each week is 248. It is known that the distribution approximates to be normal with a standard deviation of 26.
If this pattern of arrival continues, what percentage of weeks can be expected to have number of arrivals of:
(i) less than 229 per week? (ii) more than 280 per week?
(b) Pipes for tobacco are being packed in fancy plastic boxes. The length of the pipe is normally distributed with $\mu = 5''$ and $\sigma = 0.1''$. The internal length of the boxes is 5.2″. What is the probability that the box would be small for the pipe?
[Given that : $\phi(1.8) = 0.9641$, $\phi(2) = 0.9772$, $\phi(2.5) = 0.9938$]
(c) A manufacturer of envelopes knows that the weight of the envelopes is normally distributed with mean 1.9 gm and variance 0.01 square gm. Find how many envelopes weighing
(i) 2 gm or more
(ii) 2.1 gm or more, can be expected in a given packet of 1000 envelopes? (U.P.T.U. 2015)
[Given: if t is the normal variable, then $\phi(0 \le t \le 1) = 0.3413$ and $\phi(0 \le t \le 2) = 0.4772$]

Scanned with CamScanner

9. (*a*) East-East Airlines has the policy of employing only Indian woman whose height is between 62 inches and 69 inches. If the height of Indian women is approximately normally distributed with a mean of 64 inches and a standard deviation of 3 inches, Out of the 1000 applications received find the number of applicants that would be

(*i*) too tall (*ii*) too short (*iii*) of acceptable height.

(*b*) The mean height of 500 students is 151 cm and the standard deviation is 15 cm. Assuming that the heights are normally distributed, find how many students have heights between 120 and 155 cm?

(*c*) The monthly mess bill of a student who is staying in the hostel follows a normal distribution with a mean of ₹ 2000 and a standard deviation of ₹ 185. What is the probability that in the next month, his bill will go above ₹ 2400? [*U.P.T.U. (MBA) 2009*]

10. In an examination taken by 500 candidates, the average and the standard deviation of marks obtained (normally distributed) are 40% and 10%. Find approximately

(*i*) how many will pass, if 50% is fixed as a minimum?

(*ii*) what should be the minimum if 350 candidates are to pass?

(*iii*) how many have scored marks above 60%?

11. The income distribution of workers in a certain factory was found to be normal with mean ₹ 500 and standard deviation equal to ₹ 50. There were 228 persons getting above ₹ 600. How many workers were there in all?

(Area under the standard normal curve between height at 0 and 2 is 0.4772).

12. In a certain examination, the percentage of passes and distinction were 46 and 9 respectively. Estimate the average marks obtained by the candidates, the minimum pass and distinction marks being 40 and 75 respectively. Assume the distribution of marks to be normal.

$\left[\textbf{Hint. } \dfrac{40-\mu}{\sigma} = 0.1 \quad \text{and} \quad \dfrac{75-\mu}{\sigma} = 1.34.\right]$

13. The marks obtained by a no. of students for a certain subject are assumed to be approximately normally distributed with mean value 65 and S.D. of 5. If 3 students are taken at random from this set, what is the probability that exactly 2 of them will have marks over 70?

14. In an examination, it is laid down that a student passes if he secures 30% or more marks. He is placed in the first, second or third division according as he secures 60% or more marks, between 45% and 60% marks and marks between 30% and 45% respectively. He gets distinction incase he secures 80% or more marks. It is noticed from the result that 10% of the students failed in the examination, whereas 5% of them obtained distinction.

Calculate the percentage of students placed in the second division. (Assume normal distribution of marks).

[**Hint.** $P(X < 30) = 0.10$, $P(X \geq 80) = 0.05$]

15. Determine the minimum marks a student must get in order to receive an A grade if the top 10% of the students are awarded A grades in an examination where the mean mark is 72 and standard deviation is 9.

16. (*i*) When the mean of marks was 50% and S.D. 5% then 60% of the students failed in an examination. Determine the 'grace' marks to be awarded in order to show that 70% of the students passed. Assume that the marks are normally distributed.

$\left[\textbf{Hint. } \dfrac{x_1-0.5}{0.05} = 0.25, \quad \dfrac{x_2-0.5}{0.05} = -0.52\right]$

(*ii*) In a university examination of a particular year, 60% of the students failed when mean of the marks was 50% and S.D. 5%. University decided to relax the conditions of passing by lowering the pass marks to show its result 70%. Find the minimum marks for a student to pass supposing the marks to be normally distributed and no change in the performance of students takes place.

Scanned with CamScanner

17. How does a normal distribution differ from a binomial distribution? What are the important properties of normal distribution? [M.T.U. (MBA) 2012]

18. If the skulls are classified as A, B and C according as the length-breadth index is under 75, between 75 and 80 or over 80, find approximately (assuming that the distribution is normal) the mean and standard deviation of a series in which A are 58%, B are 38% and C are 4%, being given

that if $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-(x^2/2)}\, dx$ then $f(0.20) = 0.08$ and $f(1.75) = 0.46$.

[**Hint:** $P(X < 75) = 0.58$, $P(X > 80) = 0.04$]

19. The following table gives frequencies of occurrence of a variable X between certain limits:

Variable X	*Frequency*
Less than 40	30
40 or more but less than 50	33
50 and more	37

The distribution is exactly normal. Find the distribution and also obtain the fequency between $X = 50$ and $X = 60$.

20. The marks X obtained in Mathematics by 1000 students are normally distributed with mean 78% and standard deviation 11%.

Determine:

(*i*) how many students got marks above 90%?

(*ii*) What was the highest marks obtained by the lowest 10% of students?

(*iii*) Within what limits did the middle 90% of the students lie?

Answers

1. (*i*) 67 (*ii*) 134 (*iii*) 1909 **2.** (*i*) 227 (*ii*) 465
3. (*i*) $\bar{x} = 50.3, \sigma = 10.33$ (*ii*) $\mu = 65, \sigma = 10$
4. (*i*) 0.8849 (*ii*) 0.8849 (*iii*) 0.7881
5. (*i*) 143 (*ii*) 143 (*iii*) 614
6. (*a*) (*i*) 50%, (*ii*) 21.2%, (*iii*) 84% (*b*) (*i*) 1781, (*ii*) 1576, (*iii*) 1 (*c*) (*i*) 220, (*ii*) 69
7. (*a*) (*i*) 371, (*ii*) 383, (*iii*) 72.72 (*b*) (*i*) 1723 (*ii*) 1906 (*c*) (*i*) 155, (*ii*) 81, (*iii*) 23
8. (*a*) (*i*) 23% (*ii*) 11% (*b*) 0.0228 (*c*) (*i*) 159 (*ii*) 23
9. (*a*) (*i*) 48 (*ii*) 251 (*iii*) 701 (*b*) 294 (*c*) 0.0154
10. (*i*) 79 (*ii*) 35% (*iii*) 11 **11.** 10,000 **12.** 37.2
13. 0.06357 **14.** 34% **15.** 84 marks
16. (*i*) 3.85 (*ii*) 47.4. **18.** $\mu = 74.35, \sigma = 3.23$ **19.** $\mu = 46.12, \sigma = 11.76$, 25
20. (*i*) 138 (*ii*) 63.92% (*iii*) between 60 and 96.

ASSIGNMENT-IV

(For Section–A)

1. Define a random variable. Define expectation.
2. Two events A and B have probabilities 0.25 and 0.50 respectively. The probability that both events A and B occurs in 0.14. Find the probability that neither A nor B occurs. (*M.T.U.* 2013)

   [**Hint.** $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$]
3. Define probability density function and cumulative distribution function.

Scanned with CamScanner

4. What is the total probability theorem?
5. Define marginal and conditional distribution. (*M.T.U. 2012*)
6. Find the moment generating function of Poisson distribution. (*M.T.U. 2013*)
7. Find the parameters p and q of the Binomial distribution whose mean is 9 and variance is $\frac{9}{4}$. (*M.T.U. 2012*)
8. If the sum of the mean and variance of a Binomial distribution of 5 trials is $\frac{9}{5}$, find $P(X \geq 1)$. (*M.T.U. 2013*)
9. It has been found that 2% of the tools produced by a certain machine are defective. What is the probability that in a shipment of 400 such tools, 3 or more will be defective? (*M.T.U. 2013*)
10. If $P(X = 0) = P(X = 1) = k$ in a Poisson distribution then what is k?
11. For a Poisson variate X if $P(X = 1) = P(X = 2)$ then find $P(X = 4)$.
12. Find the total area under the curve of *p.d.f.* of a normal curve.
13. If for a Poisson distribution, $P(2) = P(3)$ then what is its probability function?
14. Find the parameters of a binomial distribution with mean = 8 and variance = 4.
15. If X is a normal variate with mean 30 and S.D. 5, find the probabilities that
    (*i*) $26 \leq X \leq 40$ (*ii*) $X \geq 45$ and (*iii*) $|X - 30| > 5$
16. Define :
    (*i*) Binomial distribution (*ii*) Poisson distribution (*iii*) Normal distribution
    (*A.K.T.U. 2018*)
17. Find mean, variance and third moment about mean for the binomial distribution. (*A.K.T.U. 2018*)
18. Find mean and variance of poisson distribution. (*A.K.T.U. 2018*)

Answers

2. 0.39	**6.** $M_x(t) = e^{\lambda(e^t - 1)}$	**7.** $q = \frac{1}{4}, p = \frac{3}{4}$	**8.** 0.67232
9. 0.9862	**10.** $\frac{1}{e}$	**11.** $\frac{2}{3e^2}$	**12.** 1
13. $\frac{e^{-3}(3)^x}{x!}$	**14.** $n = 16, p = \frac{1}{2}$		
15. (*i*) 0.7653	(*ii*) 0.00135	(*iii*) 0.3174.	**18.** $\lambda = np$

Scanned with CamScanner