CHAPTER

61

PROBABILITY

61.1 PROBABILITY

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

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

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

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

If the probability of the happening $= p$

and the probability of not happening $= q$

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

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

61.2 DEFINITIONS

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

   Example. The event of getting head on first coin and the event of getting tail on the second coin in a simultaneous throw of two coins are independent.

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

    (i) Odds in favour of the event = $\dfrac{m}{n}$, Odds against the event = $\dfrac{n}{m}$.
13. **Classical Definition of Probability.** If there are N equally likely, mutually, exclusive and exhaustive of events of an experiment and m of these are favourable, then the probability of the happening of the event is defined as $\dfrac{m}{N}$.
14. **Expected value.** If $p_1, p_2, p_3, \ldots, p_n$ of the probabilities of the events $x_1, x_2, x_3 \ldots x_n$ respectively the expected value

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

Example 1. *Find the probability of throwing*
(a) 5, (b) an even number with an ordinary six faced die.

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

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

(b) Total number of ways of throwing a die = 6
Number of ways falling 2, 4, 6 = 3

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

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

Solution. Total number of possible ways of throwing two dice
$= 6 \times 6 = 36$

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

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

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

Solution. A king can be chosen in 4 ways.
But a card can be drawn in 52 ways.

$\therefore$ the required probability $= \dfrac{4}{52} = \dfrac{1}{13}$ **Ans.**

EXERCISE 61.1

1. In a class of 12 students, 5 are boys and the rest are girls. Find the probability that a student selected will be a girl. **Ans.** $\dfrac{7}{12}$

2. A bag contains 7 red and 8 black balls. Find the probability of drawing a red ball. **Ans.** $\dfrac{7}{15}$

3. Three of the six vertices of a regular hexagon are chosen at random. Find the probability that the triangle with three vertices is equilateral.

Ans. $\dfrac{1}{10}$

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

(*A.M.I.E., Dec. 2009, Summer 2001*) **Ans.** $\dfrac{2}{7}$

Fill in the blanks with appropriate correct answer

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

(*A.M.I.E., Summer 2000*) **Ans.** $\dfrac{671}{1296}$

6. A pair of fair dice is thrown and one die shows a four. The probability that the other die shows 5 is ....

(*A.M.I.E., Summer 2000*) **Ans.** $\dfrac{1}{36}$

61.3 ADDITION LAW OF PROBABILITY

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

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

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

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

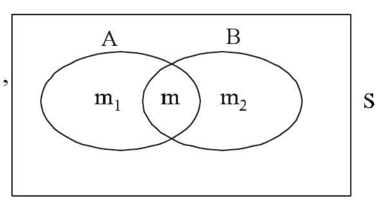
The favourable outcomes to the event A only $= m_1 - m$

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

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

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

$$A \cup B = m_1 - m) + (m_2 - m) + m$$
$$= m_1 + m_2 - \mathrm{m}$$



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

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

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

(*AMIETE, June 2010*)

Note: Mutually Exclusive Events

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

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

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

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

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

$\therefore$ Probability of drawing two balls of the same colour

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

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

Solution. There are two mutually exclusive cases,
(*i*) when the first bag is chosen, (*ii*) when the second bag is chosen.

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

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

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

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

61.4 MULTIPLICATION LAW OF PROBABILITY

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

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

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

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

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

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

Now there are four possibilities

A and B both may happen, then the number of ways $= m_1 . m_2$.
A may happen and B may fail, then the number of ways $= m_1 . n_2$
A may fail and B may happen, then the number of ways $= n_1 . m_2$
A and B both may fail, then the number of ways $= n_1 . n_2$

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

Hence the probabilities of the happening of both A and B

$$P(AB)=\frac{m_1m_2}{(m_1+n_1)(m_2+n_2)}=\frac{m_1}{m_1+n_1}.\frac{m_2}{m_2+n_2}=P(A).P(B)$$ **Proved.**

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

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

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

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

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

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

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

Example 7. *The probability that machine A will be performing an usual function in 5 years' time is $\frac{1}{4}$, while the probability that machine B will still be operating usefully at the end of the same period, is $\frac{1}{3}$*

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

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

(*ii*) *Neither will be operating.*

(*iii*) *Only machine B will be operating.*

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

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

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

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

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

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

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

= 1 – P (none of them operates)

$$=1-\frac{1}{2}=\frac{1}{2}$$ **Ans.**

Example 8. *There are two groups of subjects one of which consists of 5 science and 3 engineering subjects and the other consists of 3 science and 5 engineering subjects. An unbiased* *die is cast. If number 3 or number 5 turns up, a subject is selected at random from the first group, otherwise the subject is selected at random from the second group. Find the probability that an engineering subject is selected ultimately.*

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

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

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

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

$$= \left(\frac{1}{3}\right) \times \left(\frac{3}{8}\right) = \frac{1}{8} \qquad \text{... (1)}$$

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

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

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

$$= \frac{2}{3} \times \frac{5}{8} = \frac{5}{12} \qquad \text{... (2)}$$

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

$$= \frac{13}{24}$$ **Ans.**

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

(*i*) *the three balls are of different colours?*

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

Solution.

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

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

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

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

P (3 Blue balls or 3 Black balls) = P (3 Blue balls) + P (3 Black balls)

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

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

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

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

Required probability of drawing first white ball and second black ball

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Solution. Here there are three types of families

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

(*ii*)

Boy	Girl
0	1
1	0

Probability of zero boy in case II = 0.3 × 0.5 = 0.15

(*iii*)

Boy	Girl
0	2
1	1
2	0

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

Considering all the three cases, the probability of zero boy

$$= 0.2 + 0.15 + 0.167 = 0.517$$ **Ans.**

Example 14. *A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is* $\frac{1}{7}$ *and that of wife's selection is* $\frac{1}{5}$. *What is the probability that*

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

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

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

(*ii*) P (only one selected) = P (only husband's selection) + P (only wife's selection)\

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

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

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

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

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

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

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

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

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

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

Example 16. *A student takes his examination in four subjects* α, β, γ, δ. *He estimates his chances of passing in* α *as* $\frac{4}{5}$, *in* β *as* $\frac{3}{4}$, *in* γ *as* $\frac{5}{6}$ *and in* δ *as* $\frac{2}{3}$. *To qualify, he must pass in* α *and at least two other subjects. What is the probability that he qualifies?* (*AMIETE, Dec. 2010*)

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

There are four possibilities of passing at least two subjects.

(*i*) Probability of passing β, γ and failing δ $= \frac{3}{4} \times \frac{5}{6} \times \left(1 - \frac{2}{3}\right) = \frac{3}{4} \times \frac{5}{6} \times \frac{1}{3} = \frac{5}{24}$

(*ii*) Probability of passing γ, δ and failing β $= \frac{5}{6} \times \frac{2}{3} \times \left(1 - \frac{3}{4}\right) = \frac{5}{6} \times \frac{2}{3} \times \frac{1}{4} = \frac{5}{36}$

(*iii*) Probability of passing δ, β and failing γ $= \frac{2}{3} \times \frac{3}{4} \times \left(1 - \frac{5}{6}\right) = \frac{2}{3} \times \frac{3}{4} \times \frac{1}{6} = \frac{1}{12}$

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

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

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

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

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

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

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

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

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

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

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

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

Probability that 2 shots hit the target

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

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

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

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

= Probability of 2 shots + probability of 3 shots hitting the target

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

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

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

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

The number of combinations is 3.

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

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

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

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

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

A's total probability of winning = $\dfrac{1}{12} + \left(\dfrac{11}{12}\right)^2 \cdot \dfrac{1}{12} + \left(\dfrac{11}{12}\right)^4 \cdot \dfrac{1}{12} + \ldots$

$$= \frac{\frac{1}{12}}{1 - \left(\frac{11}{12}\right)^2} = \frac{12}{23} \qquad \left[\text{This is infinte G.P. Its sum} = \frac{a}{1-r}\right]$$

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

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

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

Hence A's chance to B's chance = $\dfrac{12}{23} : \dfrac{11}{23} = 12 : 11$ **Proved.**

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

Solution. Number of ways of throwing 6

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

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

Number of ways of throwing 7

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

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

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

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

Probability of A's winning = $p_1 + q_1 q_2 p_1 + q_1^2 q_2^2 p_1 + \ldots.$

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

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

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

EXERCISE 61.2

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

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

2. An urn contains 13 balls numbering 1 to 13. Find the probability that a ball selected at random is a ball with number that is a multiple of 3 or 4. **Ans.** $\frac{6}{13}$

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

4. A five digit number is formed by using the digits 0, 1, 2, 3, 4 and 5 without repetition. Find the probability that the number is divisible by 6. **Ans.** $\frac{4}{25}$

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

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

6. An anti-aircraft gun can take a maximum of four shots on enemy's plane moving from it. The probabilities of hitting the plane at first, second, third and fourth shots are 0.4, 0.3, 0.2 and 0.1 respectively. Find the probability that the gun hits the plane. **Ans.** 0.6976.

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

Ans. 0.000298

8. The face cards are removed from a full pack. Out of the remaining 40 cards, 4 are drawn at random. What is the probability that they belong to different suits? **Ans.** $\frac{1000}{9139}$

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

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

10. A machine contains a component *C* that is vital to its operation. The reliability of component *C* is 80%. To improve the reliability of a machine, a similar component is used in parallel to form a system S. The machine will work provided that one of these components functions correctly. Calculate the reliability of the system *S*. **Ans.** 96%

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

Ans. $\dfrac{209}{343}$

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

Ans. 0.5263

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

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

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

Ans. $\dfrac{15}{16}$

15. Tick √ the correct answer :

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

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

(*A.M.I.E.,winter 2000*) **Ans.** (*a*)

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

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

(*A.M.I.E.T.E., Summer 2000*) **Ans.** (*a*)

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

(*a*) 0.8750 (*b*) 0.0875 (*c*) 0.0625 (*d*) 0.0250

(*A.M.I.E.T.E., Summer 2001*) **Ans.** (*b*)

Fill up the blanks:

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

(*A.M.I.E.T.E, Winter 2001*) **Ans.** 0, 1

61.5 CONDITIONAL PROBABILITY

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

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \qquad \text{... (1)}$$

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

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

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

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

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

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

61.6 BAYE'S THEOREM

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

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

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

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

The events $B_1, B_2, \ldots, B_n$ being exhaustive

$$S = B_1 \cup B_2 \cup \ldots. \cup B_n$$

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

$$\therefore \qquad A = A \cap S = A \cap (B_1 \cup B_2 \cup \ldots. \cup B_n)$$

$$= (A \cap B_1) \cup (A \cap B_2) \cup \ldots. \cup (A \cap B_n) \qquad \text{[Distributive Law]}$$

$$\Rightarrow \qquad P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots. + P(A \cap B_n)$$

$$= P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + \ldots. + P(B_n)P(A/B_n)$$

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

Now, $\qquad P(A \cap B_i) = P(A)P(B_i / A)$

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

Note. $P(B)$ is the probability of occurrence B. If we are told that the event A has already occurred. On knowing about the event A, $P(B)$ is changed to $P(B/A)$. With the help of Baye's theorem we can calculate $P(B/A)$.

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

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

U_2 : the ball is drawn from urn II

W : the ball is white.

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

By Baye's Theorem

$$P(U_1/W) = \frac{P(U_1)P(W/U_1)}{P(U_1)P(W/U_1) + P(U_2)P(W/U_2)} \qquad \text{... (1)}$$

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

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

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

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

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

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

U_2 : the ball is drawn from U_2.

U_3 : the ball is drawn from U_3.

R : the ball is red.

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

By Baye's Theorem,

$$P(U_1/R) = \frac{P(U_1)P(R/U_1)}{P(U_1)P(R/U_1) + P(U_2)P(R/U_2) + P(U_3)P(R/U_3)} \qquad \text{... (1)}$$

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

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

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

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

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

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

Solution. A : bolt is manufactured by machine A.

B : bolt is manufactured by machine B.

C : bolt is manufactured by machine C.

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

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

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

By Baye's theorem

$$P(B/D) = \frac{P(B)P(D/B)}{P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C)}$$

$$= \frac{0.35 \times 0.04}{0.25 \times 0.05 + 0.35 \times 0.04 + 0.40 \times 0.02} = 0.41 \qquad \textbf{Ans.}$$

Example 24. *An insurance company insured 2000 scooter drivers 4000 car drivers and 6000 truck drivers. The probability of accidents are 0.01, 0.03 and 0.15 respectively. One of the insured persons meets with an accident. What is the probability that he is a scooter driver?* (*AMIETE, Dec. 2009*)

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

E_1 = person chosen is a scooter driver

E_2 = person chosen is a car driver

E_3 = person chosen is a truck driver, and

A = person meets with an accident.

We have,

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

Total number of persons = 2000 + 4000 + 6000 = 12000.

Therefore,

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

It is given that

$P(A/E_1)$ = Probability that a person meets with an accident given that he is a scooter driver = 0.01.

Similarly, $P(A/E_2) = 0.03$ and $P(A/E_3) = 0.15$

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

By Bayes' rule, we have

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

$$\Rightarrow \quad P(E_1/A) = \frac{\frac{1}{6} \times 0.01}{\frac{1}{6} \times 0.01 + \frac{1}{3} \times 0.03 + \frac{1}{2} \times 0.15} = \frac{1}{1+6+45} = \frac{1}{52} \qquad \textbf{Ans.}$$

EXERCISE 61.3

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

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

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

CHAPTER

62

SAMPLING METHODS

62.1 POPULATION (UNIVERSE)

Before giving the notion of sampling, we will first define *population*. The group of individuals under study is called *population* or *universe*. It may be finite or infinite.

62.2 SAMPLING

A part selected from the population is called *a sample*. The process of selection of a sample is called sampling. A *Random sample* is one in which each member of population has an equal chance of being included in it. There are ${}^{N}C_{n}$ different samples of size n that can be picked up from a population of size N.

62.3 PARAMETERS AND STATISTICS

The statistical constants of the population such as mean (μ), standard deviation (σ) are called parameters. Parameters are denoted by Greek letters.

The mean $(\bar{x})$, standard deviation $|S|$ of a sample are known as statistics. Statistics are denoted by Roman letters.

Symbols for Population and Samples

Characteristic	Population	Sample
	Parameter	Statistic
Symbols	population size = N population mean = μ population standard deviation = σ population proportion = p	sample size = n sample mean = $\bar{x}$ sample standard deviation = s sample proportion = $\tilde{p}$

62.4 AIMS OF A SAMPLE

The population parameters are not known generally. Then the sample characteristics are utilised to approximately determine or estimate of the population. Thus, static is an estimate of the parameter. To what extent can we depend on the sample estimates?

The estimate of mean and standard deviation of the population is a primary purpose of all scientific experimentation. The logic of the sampling theory is the logic of *induction.* In induction, we pass from a particular (sample) to general (population). This type of generalization here is known as *statistical inference.* The conclusion in the sampling studies are based not on certainties but on probabilities.

62.5 TYPES OF SAMPLING

Following types of sampling are common:

(1) Purposive sampling (2) Random sampling (3) Stratified sampling (4) Systematic sampling

62.6 SAMPLING DISTRIBUTION

From a population a number of samples are drawn of equal size n. Find out the mean of each sample. The means of samples are not equal. The means with their respective frequencies are grouped. The frequency distribution so formed is known as *sampling distribution of the mean*. Similarly, sampling distribution of standard deviation we can have.

62.7 STANDARD ERROR (S.E.)

is the standard deviation of the sampling distribution. For assessing the difference between the expected value and observed value, standard error is used. Reciprocal of standard error is known as *precision.*

62.8 SAMPLING DISTRIBUTION OF MEANS FROM INFINITE POPULATION

Let the population be infinitely large and having a population mean of μ and a population variance of σ^2. If x is a random variable denoting the measurement of the characteristic, then

Expected value of x, $E(x) = \mu$

Variance of x, Var $(x) = \sigma^2$

The sample mean $\bar{x}$ is the sum of n random variables, *viz.*, $x_1, x_2, \ldots, x_n$, each being divided by n. Here, $x_1, x_2, \ldots, x_n$ are independent random variables from the infinitely large population.

$\therefore$ $E(x_1) = \mu$ and $Var(x_1) = \sigma^2$

$E(x_2) = \mu$ and $Var(x_2) = \sigma^2$ and so on

Finally $E(\bar{x}) = E\left[\dfrac{x_1 + x_2 + \ldots + x_n}{n}\right] = \dfrac{1}{n}Ex_1 + \dfrac{1}{n}E(x_2) + \ldots + \dfrac{1}{n}E(x_n) = \dfrac{1}{n}\mu + \dfrac{1}{n}\mu + \ldots + \dfrac{1}{n}\mu = \mu$

and $Var(\bar{x}) = Var\left[\dfrac{x_1 + x_2 + \ldots + x_n}{n}\right] = Var\left(\dfrac{x_1}{n}\right) + Var\left(\dfrac{x_2}{n}\right) + \ldots + Var\left(\dfrac{x_n}{n}\right)$

$$= \frac{1}{n^2}Var(x_1) + \frac{1}{n^2}Var(x_2) + \ldots + \frac{1}{n^2}Var(x_n) = \frac{1}{n^2}.\sigma^2 + \frac{1}{n^2}.\sigma^2 + \ldots + \frac{1}{n^2}.\sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

The expected value of the sample mean is the same as population mean. The variance of the sample mean is the variance of the population divided by the sample size.

The average value of the sample tends to true population mean. If sample size *(n)* is increased then variance of $\bar{x}$, $\left(\dfrac{\sigma^2}{n}\right)$ gets reduced, by taking large value of n, the variance $\left(\dfrac{\sigma^2}{n}\right)$ of $\bar{x}$ can be made as small as desired. The standard deviation $\left(\dfrac{\sigma}{\sqrt{n}}\right)$ of $\bar{x}$ is also called **standard error of the mean.** It is denoted by $\sigma_{\bar{x}}$.

Sampling with Replacement

When the sampling is done with replacement, so that the population is back to the same form before the next sample member is picked up. We have

$$E(\bar{x}) = \mu$$

$$Var(\bar{x}) = \frac{\sigma^2}{n} \text{ or } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Sampling without replacement from Finite population

When a sample is picked up without replacement from a finite population, the probability distribution of second random variable depends on the outcome of the first pick up. n sample members do not remain independent. Now we have

$$E(\bar{x}) = \mu$$

and $$Var(\bar{x}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} \quad \text{or} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$= \frac{\sigma}{\sqrt{n}} \text{ app} \qquad \left(\text{if } \frac{n}{N} \text{ is very small}\right)$$

Sampling from Normal Population

If $x \sim N\left(\mu, \sigma^2\right)$ then it follows that $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

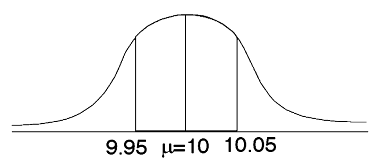
Example 1. *The diameter of a component produced on a semi-automatic machine is known to be distributed normally with a mean of 10 mm and a standard deviation of* 0.1 *mm. If we pick up a random sample of size 5, what is the probability that the same mean will be between 9.95 and 10.05 mm?*

Solution. Let x be a random variable representing the diameter of one component picked up at random.

Here $x \sim$ N (10, 0.01), Therefore, $\bar{x} \sim N\left(10, \frac{0.01}{5}\right)$ $\qquad \left[\bar{x} = N\left(\bar{x}, \frac{\sigma^2}{n}\right)\right]$

$$Pr\{9.95 \le \bar{x} \le 10.05\} = 2 \times Pr\{10 \le \bar{x} \le 10.05\} \qquad \left\{z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right\}$$

$$= 2 \times Pr\left\{\frac{10-\mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{10.05-\mu}{\frac{\sigma}{\sqrt{n}}}\right\}$$



$$= 2 \times Pr\left\{0 \le z \le \frac{10.05-10}{\frac{0.1}{\sqrt{5}}}\right\} = 2 \times Pr\{0 \le z \le 1.12\} = 2 \times 0.3686 = 0.7372 \qquad \textbf{Ans.}$$

Similar Question

A sample of size 25 is picked up at random from a population which is normally distributed with a mean 100 and a variance of 36. Calculate (*a*) $Pr\{\bar{x} \le 99\}$, (*b*) $Pr\{98 \le \bar{x} \le 100\}$

Ans. (*a*) 0.2023 (*b*) 0.4522

62.9 SAMPLING DISTRIBUTION OF THE VARIANCE

We use a sample statistic called the sample variance to estimate the population variance. The sample variance is usually denoted by s^2

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

62.10 TESTING A HYPOTHESIS

On the basis of sample information, we make certain decisions about the population. In taking such decisions we make certain assumptions. These assumptions are known as *statistical hypothesis.* These hypothesis are tested. Assuming the hypothesis correct we calculate the probability of getting the observed sample. If this probability is less than a certain assigned value, the hypothesis is to be rejected.

62.11 NULL HYPOTHESIS (H_0)

Null hypothesis is based for analysing the problem. Null hypothesis is the *hypothesis of no difference*. Thus, we shall persume that there is no significant difference between the observed value and expected value. Then, we shall test whether this hypothesis is satisfied by the data or not. If the hypothesis is not approved the difference is considered to be significant. If hypothesis is approved then the difference would be described as due to sampling fluctuation. Null hypothesis is denoted by H_0.

62.12 ERRORS

In sampling theory to draw valid inferences about the population parameter on the basis of the sample results.

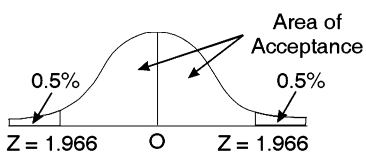
We decide to accept or to reject the lot after examining a sample from it. As such, we are liable to commit the following two types of errors.

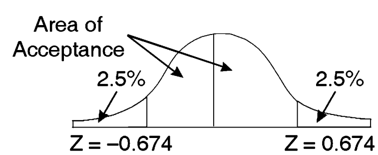
Type 1 Error. If H_0 is rejected while it should have been accepted.

Type II Error. If H_0 is accepted while it should have been rejected.

62.13 LEVEL OF SIGNIFICANCE

There are two critical regions which cover 5% and 1% areas of the normal curve. The shaded portions are the critical regions.





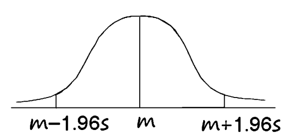
Thus, the probability of the value of the variate falling in the critical region is the level of significance. If the variate falls in the critical area, the hypothesis is to be rejected.

62.14 TEST OF SIGNIFICANCE

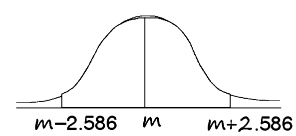
The tests which enables us to decide whether to accept or to reject the null hypothesis is called the tests of significance. If the difference between the sample values and the population values are so large (lies in critical area), it is to be rejected

62.15 CONFIDENCE LIMITS

$\mu - 1.96\ \sigma, \mu + 1.96\sigma$ are 95% confidence limits as the area between $\mu - 1.96\sigma$ and $\mu + 1.96\sigma$ is 95%. If a sample statistics lies in the interval $\mu - 1.96\ \sigma$, $\mu + 1.96\sigma$, we call 95% confidence interval.



Similarly, $\mu - 2.58\sigma$, $\mu + 2.58\sigma$ is 99% confidence limits as the area between $\mu - 2.58\sigma$ and $\mu + 2.58\sigma$ is 99%. The numbers 1.96, 2.58 are called confidence coefficients.



62.16 TEST OF SIGNIFICANCE OF LARGE SAMPLES *(N* > 30)

Normal distribution is the limiting case of Binomial distribution when n is large enough. For normal distribution 5% of the items lie outside $\mu \pm 1.96\sigma$ while only 1% of the items lie outside $\mu \pm 2.586\ \sigma$.

$$z = \frac{x - \mu}{\sigma}$$

where z is the standard normal variate and x is the observed number of successes.

First we find the value of z. Test of significance depends upon the value of z.

(*i*) (*a*) If $|z| < 1.96$, difference between the observed and expected number of successes is not significant at the 5% level of significance.

(*b*) If $|z| > 1.96$, difference is significant at 5% level of significance.

(*ii*) (*a*) If $|z| < 2.58$, difference between the observed and expected number of successes is not significant at 1% level of significance.

(*b*) If $|z| > 2.58$, difference is significant at 1% level of significance.

Example 2. *A cubical die was thrown 9,000 times and 1 or 6 was obtained 3120 times. Can the deviation from expected value lie due to fluctuations of sampling?*

Solution. Let us consider the hypothesis that the die is an unbiased one and hence the probability of obtaining 1 or $6 = \frac{2}{6} = \frac{1}{3}$ *i.e.*, $p = \frac{1}{3},\ q = \frac{2}{3}$

The expected value of the number of successes $= np = 9000 \times \frac{1}{3} = 3000$

Also $\sigma = \text{S.D.} = \sqrt{npq} = \sqrt{9000 \times \frac{1}{3} \times \frac{2}{3}} = \sqrt{2000} = 44.72$

$$3\,\sigma = 3 \times 44.72 = 134.16$$

Actual number of successes = 3120

Difference between the actual number of successes and expected number of successes

$$= 3120 - 3000 = 120 \text{ which is } < 3\,\sigma$$

Hence, the hypothesis is correct and the deviation is due to fluctuations of sampling due to randon causes. **Ans.**

62.17 SAMPLING DISTRIBUTION OF THE PROPORTION

A simple sample of n items is drawn from the population. It is same as a series of n independent trials with the probability p of success. The probabilities of 0, 1, 2, ..., n success are the terms in the binomial expansion of $(q + p)^n$.

Here mean $= np$ and standard deviation $= \sqrt{npq}$.

Let us consider the proportion of successes, then

(*a*) Mean proportion of successes $= \frac{np}{n} = p$

(*b*) Standard deviation (standard error) of proportion of successes $= \frac{\sqrt{npq}}{n} = \sqrt{\frac{pq}{n}}$

(*c*) Precision of the proportion of success $= \frac{1}{\text{S.E.}} = \sqrt{\frac{n}{pq}}$.

Example 3. *A group of scientific mens reported 1705 sons and 1527 daughters. Do these figures conform to the hypothesis that the sex ratio is $\frac{1}{2}$.*

Solution. The total number of observations = 1705 + 1527 = 3232

The number of sons = 1705

Therefore, the observed male ratio $= \frac{1705}{3232} = 0.5275$

On the given hypothesis the male ratio = 0.5000

Thus, the difference between the observed ratio and theoretical ratio

$$= 0.5275 - 0.5000$$
$$= 0.0275$$

The standard deviation of the proportion $= s = \sqrt{\dfrac{pq}{n}} = \sqrt{\dfrac{\frac{1}{2}\times\frac{1}{2}}{3232}} = 0.0088$

The difference is more than 3 times of standard deviation.

Hence, it can be definitely said that the figures given do not conform to the given hypothesis.

62.18 ESTIMATION OF THE PARAMETERS OF THE POPULATION

The mean, standard deviation etc. of the population are known as parameters. They are denoted by μ and σ. Their estimates are based on the sample values. The mean and standard deviation of a sample are denoted by $\bar{x}$ and s respectively. Thus, a static is an estimate of the parameter. There are two types of estimates.

(*i*) *Point estimation:* An estimate of a population parameter given by a single number is called a point estimation of the parameter. For example,

$$(\text{S.D.})^2 = \frac{\sum (x-\bar{x})^2}{n-1}$$

(*ii*) *Interval estimation:* An interval in which population parameter may be expected to lie with a given degree of confidence. The intervals are

(*i*) $\bar{x} - \sigma_s$ to $\bar{x} + \sigma_s$ (68.27% confidence level)

(*ii*) $\bar{x} - 2\sigma_s$ to $\bar{x} + 2\sigma_s$ (95.45% confidence level)

(*iii*) $\bar{x} - 3\sigma_s$ to $\bar{x} + 3\sigma_s$ (99.73% confidence level)

$\bar{x}$ and σ_s are the mean and S.D. of the sample.

Similarly, $\bar{x} \pm 1.96\,\sigma_s$ and $\bar{x} \pm 2.58\,\sigma_s$ are 95% and 99% confidence of limits for μ.

$\bar{x} \pm 1.96\dfrac{\sigma}{\sqrt{n}}$ and $\bar{x} \pm 2.58\dfrac{\sigma}{\sqrt{n}}$ are also the intervals as $\sigma_s = \dfrac{\sigma}{\sqrt{n}}$.

62.19 COMPARISON OF LARGE SAMPLES

Let two large samples of size n_1, n_2 be drawn from two populations of proportions of attributes A's as P_1, P_2 respectively.

(*i*) *Hypothesis:* As regards the attribute A, the two populations are similar. On combining the two samples

$$P = \frac{n_1P_1 + n_2P_2}{n_1 + n_2}$$

where p is the common proportion of attributes.

Let e_1, e_2 be the standard errors in the two samples, then

$$e_1^2 = \frac{pq}{n_1} \text{ and } e_2^2 = \frac{pq}{n_2}$$

If e be the standard error of the combined samples, then

$$e = P_1^2 + P_2^2 = \frac{pq}{n_1} + \frac{pq}{n_2} = pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right]$$

$$z = \frac{P_1 - P_2}{e}$$

1. If $z > 3$, the difference between P_1 and P_2 is significant.
2. If $z < 2$, the difference may be due to fluctuations of sampling.
3. If $2 < z < 3$, the difference is significant at 5% level of significance.

(*ii*) *Hypothesis.* In the two populations, the proportions of attribute A are not the same, then standard error e of the difference $p_1 - p_2$ is

$$e^2 = p_1 + p_2$$

$$= \frac{P_1 - q_1}{n_1} + \frac{P_2 - q_2}{n_2}, z = \frac{P_1 - P_2}{e} < 3,$$

difference is due to fluctuations of samples.

Example 4. *In a sample of 600 men from a certain city, 450 are found smokers. In another sample of 900 men from another city, 450 are smokers. Do the data indicate that the cities are significantly different with respect to the habit of smoking among men.*

Solution. $n_1 = 600$ men, Number of smokers= 450, $P_1 = \frac{450}{600} = 0.75$

$n_2 = 900$ men, Number of smokers= 450, $P_2 = \frac{450}{900} = 0.5$

$$P = \frac{n_1P_1 + n_2P_2}{n_1 + n_2} = \frac{600 \times 0.75 + 900 \times 0.5}{600 + 900} = \frac{900}{1500} = 0.60$$

$$q = 1 - P = 1 - 0.6 = 0.4$$

$$e^2 = P_1^2 + P_2^2 = Pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$e^2 = 0.6 \times 0.4\left(\frac{1}{600} + \frac{1}{900}\right) = 0.000667$$

$$e = 0.02582$$

$$z = \frac{P_1 - P_2}{e} = \frac{0.75 - 0.50}{0.02582} = 9.682$$

$z > 3$ so that the difference is significant. **Ans.**

Example 5. *One type of aircraft is found to develop engine trouble in 5 flights out of a total of 100 and another type in 7 flights out of a total of 200 flights. Is there a significant difference in the two types of aircrafts so far as engine defects are concerned.*

Solution. $n_1 = 100$ flights, Number of troubled flights = 5, $P_1 = \frac{5}{100} = \frac{1}{20}$

$n_2 = 200$ flights, Number of troubled flights = 7, $P_2 = \frac{7}{200}$

$$e^2 = \frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2} = \frac{0.05 \times 0.95}{100} + \frac{0.035 \times 0.965}{200}$$

$$= 0.000475 + 0.0001689 = 0.0006439$$

$$e = 0.0254$$

$$z = \frac{0.05 - 0.035}{0.0254} = 0.59$$

$z < 1$, Difference is not significant. **Ans.**

62.20 THE t-DISTRIBUTION (FOR SMALL SAMPLE)

The students distribution is used to test the significance of

(*i*) The mean of a small sample.

(*ii*) The difference between the means of two small samples or to compare two small samples.

(*iii*) The correlation coefficient.

Let $x_1, x_2, x_3, \ldots, x_n$, be the members of random sample drawn from a normal population with mean μ. If $\bar{x}$ be the mean of the sample then

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \quad \text{where } s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$$

Example 6. *A machine which produces mica insulating washers for use in electric device to turn out washers having a thickness of 10 mm. A sample of 10 washers has an average thickness 9.52 mm with a standard deviation of 0.6 mm. Find out t.*

Solution. $\bar{x} = 9.52, M = 10, S' = 0.6, n = 10$

$$t = \frac{\bar{x} - M}{\frac{s}{\sqrt{n}}} = \frac{9.52 - 10}{\frac{0.6}{\sqrt{10}}} = -\frac{0.48\sqrt{10}}{0.6} = -\frac{4}{5}\sqrt{10}$$

$$= -0.8 \times 3.16 = -2.528$$ **Ans.**

62.21 WORKING RULE

To calculate significance of sample mean at 5% level.

Calculate $t = \frac{\bar{x} - \mu}{s}\sqrt{n}$ and compare it to the value of t with (n- 1) degrees of freedom at 5% level, obtained from the table. Let this tabulated value of t be t_1.

If $t < t_1$, then we accept the hypothesis *i.e.*, we say that the sample is drawn from the population.

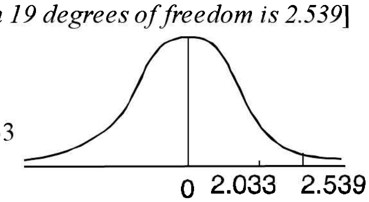
If $t > t_1$, we compare it with the tabulated value of t at 1% level of significance for (n – 1) degrees of freedom. Denote it by t_2. If $t_1 < t < t_2$ then we say that the value of t is significant.

If $t > t_1$, we reject the hypothesis and the sample is not drawn from the population.

Example 7. *A manufacturer intends that his electric bulbs have a life of 1000 hours. He tests a sample of 20 bulbs, drawn at random from a batch and discovers that* the mean life of the sample bulbs is 990 hours with a *S.D of 22 hours. Does this signify that the batch is not up to the standard?*

[**Given:** *The table value of t at 1% level is significance with 19 degrees of freedom is 2.539*]

Solution. $\bar{x} = 990, \sigma = 22, x = 1000$



$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{990 - 1000}{\frac{22}{\sqrt{20}}} = -\frac{10\sqrt{20}}{22} = -\frac{22.36}{11} = -2.033$$

Since the calculated value of t (2.032) is less than the value of t (2.539) from the table. Hence, it is not correct to say that this batch is not upto this standard. **Ans.**

Example 8. *Ten individuals are chosen at random from a population and their heights are found to be in inches 63, 63, 64, 65, 66, 69, 69, 70, 70, 71. Discuss the suggestion that the Mean height of universe is 65.*

For 9 degree of freedom t at 5% level of significance = 2.262.

Solution.

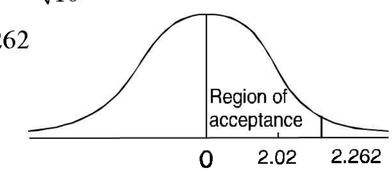
x	$x-67$	$(x-67)^2$
63	–4	16
63	–4	16
64	–3	9
65	–2	4
66	–1	1
69	+2	4
69	+2	4
70	+3	9
70	+3	9
71	+4	16
$\sum x = 670$		$\sum (x-\bar{x})^2 = 88$

$$\bar{x} = \frac{\sum x}{n} = \frac{670}{10} = 67,$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} = \sqrt{\frac{88}{9}} = 3.13$$

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{67-65}{\frac{3.13}{\sqrt{10}}} = \frac{2\sqrt{10}}{3.13} = 2.02$$

$2.02 < 2.262$



Calculated value of t (2.02) is less than the table value of t (2.262). The hypothesis is accepted the mean height of universe is 65 inches. **Ans.**

Example 9. *The mean life time of sample of 100 fluorescent light bulbs produced by a company is computed to be 1570 hours with a standard deviation of 120 hours. The company claims that the average life of the bulbs produced by it is 1600 hours. Using the level of significance of 0.05, is the claim acceptable?*

Solution. $\bar{x} = 1570$, $s = 120$, $n = 100$, $\mu = 1600$

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{1570-1600}{\frac{120}{\sqrt{100}}} = \frac{1570-1600}{12} = 2.5$$

At 0.05 the level of significance, $t = 1.96$

Calculated value of $t >$ Table value of t.

$2.5 > 1.96$

Hence the claim is to be rejected. **Ans.**

Example 10. *A sample of 6 persons in an office revealed an average daily smoking of 10, 12, 8, 9, 16, 5 cigarettes. The average level of smoking in the whole office has to be estimated at 90% level of confidence.*

$t = 2.015$ for 5 degree of freedom

Solution.

x	$x-10$	$(x-10)^2$
10	0	0
12	2	4
8	–2	4
9	–1	1
16	+6	36
5	–5	25
Total	0	$\sum (x-10)^2 = 70$

$$\text{Mean} = a + \frac{\Sigma fd}{\Sigma f} = 10 + \frac{0}{6} = 10$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} = \sqrt{\frac{70}{5}} = 3.74$$

At 90% level of confidence, $t = \pm\, 2.015$.

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} \Rightarrow \pm\, 2.015 = \frac{10-\mu}{\frac{3.74}{\sqrt{6}}}$$

$$\Rightarrow \quad \mu = 2.015 \times \frac{3.74}{\sqrt{6}} + 10 = 6.92, 13.08$$ **Ans.**

Example 11. *A fertiliser mixing machine is set to give 12 kg of nitrate for quintal bag of fertiliser: Ten 100 kg bags are examined The percentages of nitrate per bag are as follows:*

11, 14, 13, 12, 13, 12, 13, 14, 11, 12

Is there any reason to believe that the machine is defective? Value of t for 9 degrees of freedom is 2.262.

Solution. The calculation of $\bar{x}$ and s is given in the following table:

x	$d = x - 12$	d^2
11	–1	1
14	2	4
13	1	1
12	0	0
13	1	1
12	0	0
13	1	1
14	2	4
11	–1	1
12	0	0
$\sum x = 125$	$\sum d = 5$	$\sum d^2 = 13$

$$\mu = 12 \text{ kg}, \; n = 10, \bar{x} = \frac{\sum x}{n} = \frac{125}{10} = 12.5$$

$$s^2 = \frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2 = \frac{13}{10} - \left(\frac{5}{10}\right)^2 = \frac{13}{10} - \frac{1}{4} = \frac{21}{20} = \frac{105}{100}$$

$$s = 1.024$$

Value of t for 9 degrees of freedom = 2.262

$$\text{Also } t = \frac{\bar{x} - \mu}{s}\sqrt{n} = \frac{12.5 - 12}{1.024}\sqrt{10} = 1.54$$

Since the value of t is less than 2.262, there in no reason to believe that machine is defective. **Ans.**

Example 12. *A random sample of size 16 values from a normal population showed a mean of 53 and a sum of squares of deviation from the mean equals to 150. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.*

$\gamma = 15, \; \alpha = 0.05, t = 2.131$

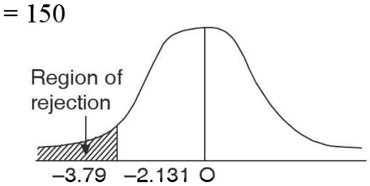
$\alpha = 0.01, t = 2.947$

Solution. $\mu = 56, n = 16, \; \bar{x} = 53, \; \sum (x - \bar{x})^2 = 150$

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{150}{15} = 10$$

$$s = \sqrt{10}$$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{53 - 56}{\frac{\sqrt{10}}{\sqrt{16}}} = \frac{-3 \times 4}{\sqrt{10}} = -3.79$$



$$t = 3.79$$

When α = 0.5 then $3.79 > 2.131$

When α = 0.01 then $3.79 > 2.947$

Thus, the sample cannot be regarded as taken from the population. **Ans.**

62.22 TESTING FOR DIFFERENCE BETWEEN MEANS OF TWO SMALL SAMPLES

Let the mean and variance of the first population be μ_1 and σ_1^2 and $\mu_2 . \ \sigma_2^2$ be the mean and variance of the second population.

Let $\bar{x}_1$ be the mean of small sample of size n_1 from first population and $\bar{x}_2$ the mean of a sample of size n_2 from second population.

We know that

$$E(\bar{x}_1) = \mu_1 \text{ and } Var\ (\bar{x}_1) = \frac{\sigma_1^2}{n_1}$$

$$E(\overline{x}_2) = \mu_2 \text{ and } Var(\overline{x}_2) = \frac{\sigma_2^{\ 2}}{n_2}$$

If the samples are independent, then $(\overline{x}_1)$ and $(\overline{x}_2)$ are also independent.

$$E(\overline{x}_1 - \overline{x}_2) = \mu_1 - \mu_2 \text{ and } Var(\overline{x}_1 - \overline{x}_2) = \frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}$$

$$\overline{x}_1 \sim N\left(\mu_1, \frac{\sigma^2_1}{n_1}\right) \text{ and } \overline{x}_2 \sim N\left(\mu_2, \frac{\sigma^2_2}{n_2}\right) \text{ then } (\overline{x}_1 - \overline{x}_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}\right)$$

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}}$$

If the population is the same then

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (\mu_1 - \mu_2 = \mu_1 - \mu_1 = 0)$$

Example 13. *Two independent samples of 8 and 7 items respectively had the following values of the variable (weight in ounces):*

Sample 1: 9 11 13 11 15 9 12 14

Sample 2: 10 12 10 14 9 8 10

Is the difference between the means of the sample significant?

[*Given for $V = 13$, $t_{0.05} = 2.16$*]

Solution.

Assumed mean of $x = 12$, Assumed mean of $y = 10$

x	$(x–12)$	$(x–12)^2$	y	$(y–10)$	$(y–10)^2$
9	–3	9	10	0	0
11	–1	1	12	2	4
13	1	1	10	0	0
11	–1	1	14	4	16
15	3	9	9	–1	1
9	–3	9	8	–2	4
12	0	0	10	0	0
14	2	4	–	–	–
94	–2	34	73	3	25

$$\overline{x} = \frac{\sum x}{n} = \frac{94}{8} = 11.75$$

$$\sigma_x^{\ 2} = \frac{\Sigma(x-12)^2}{n} - \left(\frac{\Sigma(x-12)}{n}\right)^2 = \frac{34}{8} - \left(\frac{-2}{8}\right)^2 = 4.1875$$

$$\overline{y} = \frac{\sum y}{n} = \frac{73}{7} = 10.43$$

$$\sigma_y^{\ 2} = \frac{\sum(y-10)^2}{n} - \left[\frac{\Sigma(y-10)}{n}\right]^2 = \frac{25}{7} - \left(\frac{3}{7}\right)^2 = 3.388$$

$$s = \sqrt{\frac{(x-\overline{x})^2 + \sum(y-\overline{y})^2}{n_1 + n_2 - 2}} = \sqrt{\frac{34+25}{8+7-2}} = \sqrt{\frac{59}{13}} = \sqrt{4.54} = 2.13$$

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{11.75-10.43}{2.13\sqrt{\frac{1}{8}+\frac{1}{7}}} = \frac{1.32}{2.13\sqrt{0.268}} = \frac{1.32}{2.13\times 0.518} = \frac{1.32}{1.103} = 1.2$$

Thus, 5% value of t for 13 degree of freedom is given to be 2.16. Since calculated value of t is 1.2 is less than 2.16, the difference between the means of samples is not significant. **Ans.**

EXERCISE 62.1

1. A random sample of six steel beams has mean compressive strength of 58.392 psi (pounds per square inch) with a standard deviation of s = 648 psi. Test the null hypothesis $H_0 = \mu$ = 58,000 psi against the alternative hypothesis H_1: μ > 58,000 psi at 5% level of significance (value for t at 5 degree of freedom and 5% significance level is 2.0157). Here μ denotes the population mean. *(A.M.I.E., Summer 2000)*

2. A certain cubical the was thrown 96 times and shows 2 upwards 184 times. Is the the biased?

Ans. die is biased.

3. In a sample of 100 residents of a colony 60 are found to be wheat eaters and 40 rice eaters. Can we assume that both food articles are equally popular?

4. Out of 400 children, 150 are found to be under weight. Assuming the conditions of simple sampling, estimate the percentage of children who are underweight in, and assign limits within which the per-centage probably lies.

Ans. 37.5% approx. Limits = 37.5 ± 3 (2.4)

5. 500 eggs are taken at random from a large consignment, and 50 are found to be bad. Estimate the percentage of bad eggs in the consignment and assign limits within which the percentage probably lies.

Ans. 10%, 10 ± 3.9

6. A machine puts out 16 imprefect articles in a sample of 500. After the machine is repaired, puts out 3 imprefect articles in a batch of 100. Has the machine been improved?

Ans. The machine has not been improved.

7. In a city *A*, 20% of a random sample of 900 school boys had a certain slight physical defect. In another city *B*, 18.5% of a random sample of 1600 school boys had the same defect. Is the difference between the proportions significant?

Ans. z = 0.37, Difference between proportions is significant.

8. In two large populations there are 30% and 25% respectively of fair haired people. Is this difference likely to be hidden in samples of 1200 and 900 respectively from the two populations?

Ans. z = 2.5, not hidden at 5% level of significance.

9. One thousand articles from a factory are examined and found to be three percent defective. Fifteen hundred similar articles from a second factory are found to be only 2 percent defective. Can it reasonably be concluded that the product of the first factory is inferior to the second?

Ans. It cannot be reasonable concluded that the product of the first factory is inferior to that of the second.

10. A manufacturing company claims 90% assurance that the capacitors manufactured by them will show a tolerance of better than 5%. The capacitors are packaged and sold in lots of 10. Show that about 26% of his customers ought to complain that capacitors do not reach the specified standard.

62.23 CHI SQUARE TEST

The Chi-square distribution is one of the most extensively used distribution function in statistics. It was first discovered by Helmert in 1875 and later on Karl Pearson's in 1900.

62.24 CHI-SQUARE VARIATES

The square of a standard normal variate is known as Chi-square variate (χ^2) with one degree function :

$$z = \frac{x-\mu}{\sigma} \text{ is a normal variate.}$$

Hence $\left(\frac{x-\mu}{\alpha}\right)^2$ is a Chi-square variate.

If x be a normaly distributed variate and $x_1, x_2, \ldots\ldots\ldots, x_n$ be a random sample of n-values from this population then $w = x_1^2 + x_2^2 + \ldots\ldots + x_n^2$ has χ^2 distribution with n-degree of freedom.

62.25 CONDITIONS FOR CHI-SQUARE TEST

There are some conditions which are necessary for Chi square test.

1. The sample under study must be large and may be total of cell frequency should not be less than 50.
2. The member of the cells should be independent.
3. The cell frequency of each cell should be greater than 5. If any cell has frequency less than 5 then it should be combined with the next or preceding cell until the total frequency exceeds 5.
4. If there are any constraint on the cell frequencies they should be linear
   i.e.; $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots\ldots + \alpha_i x_i + \ldots\ldots\ldots + \alpha_n x_n = \lambda$

Note: Cell frequency should not involve any logarithmic, exponential or trigonometric relation.

62.26 CHI-SQUARE (χ^2) IS USED AS:

(1) Test of independence (2) Test of goodness of fit

(3) To test if the hypothetical value of the population variate is σ^2

(4) To test the homogeneity of independent estimate of the population variance.

We shall mainly use the first two test

62.27 CHI-SQUARE TEST OF GOODNESS OF FIT

This test is used to test significance of the discrepancy between theory and experiment. It helps us to find if the deviation of the experiment from the theory is just by chance or it is due to the inadequacy of the theory to fit the observed data.

The theoretical frequencies for various classes are calculated from the assumption of the population. The significant deviation between the observed and theoretical frequencies is tested by means of this test.

χ is calculated by means of the following formula

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \quad \text{and} \quad \Sigma O_i = \Sigma E_i = N$$

where O_i is the observed frequency E_i is the expected (Theoretical) frequency of the cell.

62.28 WORKING RULE TO CALCULATE χ^2:

Step 1. Calculate the expected frequencies.

Step 2. Calculate the difference between each observed frequency O_i and the corresponding expected frequency E_i for each class i.e.; to find $O_i - E_i$

Step 3. Square the difference obtained in step 2 for each value i.e.; Calculate $(O_i - E_i)^2$.

Step 4. Divide $(O_i - E_i)^2$ by the expected frequency E to get $\dfrac{(O_i - E_i)^2}{E_i}$

Step 5. Add all these quotients obtained in step 4. Then $\chi^2 = \dfrac{\sum_{i=1}^{n} (O_i - E_i)^2}{E_i}$

It is to be noted

(1) The value of χ^2 is always positive. (2) χ^2 will be zero if each pair is zero.

(3) The value of χ^2 lies between 0 and ∞.

62.29 DEGREE OF FREEDOM

Case I. If the data is given in the form of a series of variables in a row or column then the degree of freedom = (No. of items in the series) – 1

Case 2. When the number of frequencies are put in cells in a contingency table.

The degree of freedom = (R – 1) (C – 1)

where R is number of rows and C is the number of columns.

Example 14. *A survey of 320 families with 5 children is given below :*

No. of boys	*5*	*4*	*5*	*2*	*1*	*0*	*Total*
No. of girls	*0*	*1*	*2*	*3*	*4*	*5*	
No. of families	*14*	*56*	*110*	*88*	*40*	*12*	*320*

Is this result consistent with hypothesis i.e.; the male and female **birth** *are equally possible.*

Solution. Null Hypothesis H_0.

(1) Male and Female birth are equally probable.

Alternate Hypothesis $\mathbf{H_1}$: Male and female birth are not equally probable.

Calculation of expected frequencies $(q + p)^n$

Probability of female birth $= p = \dfrac{1}{2}$

Probability of male birth $= q = \dfrac{1}{2}$

$(q + p)^n = q^n + {}^nC_1\, p\, q^{n-1} + {}^nC_2\, p^2 q^{n-2} + {}^nC_3\, p^3 q^{n-3} + \ldots\ldots + p^n$

$$\left(\frac{1}{2}+\frac{1}{2}\right)^5 = \left(\frac{1}{2}\right)^5 + 5\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 + 10\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 + 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 + 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^5$$

$$\text{No. of girls} = 320\left[\frac{1}{32}+\frac{5}{32}+\frac{10}{32}+\frac{10}{32}+\frac{5}{32}+\frac{1}{32}\right]$$

$$= 320 \times \frac{1}{32} + 320 \times \frac{5}{32} + 320 \times \frac{10}{32} + 320 \times \frac{10}{32} + 320 \times \frac{5}{32} + 320 \times \frac{1}{32}$$

$$= 10 + 50 + 100 + 100 + 50 + 10$$

These are the expected frequencies of the female births.

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
14	10	4	16	1.60
56	50	6	36	0.72
110	100	10	100	1.00
88	100	– 12	144	1.44
40	50	– 10	100	2.00
12	10	2	4	0.40
			Total	7.16

Level of significance Let $\alpha = 0.05$

Critical value. The table value of χ^2 at $\alpha = 0.05$ for $(6 - 1)(2 - 1) = 5$ degree of freedom is 11.07

Decision Since the calculated value of χ^2 (7.16) < Table value of χ^2 at level of significance 0.05 for $5df = 11.07$

Hence, the null hypothesis is accepted i.e.; the male and female birth is equally probable.

Example 15. *The table below give the number of air craft accidents that occured during the various days of the week. Test whether the accidents are uniformly distributed over the week.*

Days	*Mon.*	*Tue.*	*Wed.*	*Thu.*	*Fri*	*Sat*	*Sun*	*Total no. accidents*
No. of accidents	*14*	*18*	*12*	*11*	*15*	*14*	*14*	*98*

Solution. H_0 : Null Hypothesis : The accidents are uniformly distributed over the week.

The expected frequencies of the accidents on each day = $\frac{98}{7} = 14$

O	E	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
14	14	0	0	0
18	14	4	16	1.14
12	14	– 2	4	0.29
11	14	– 3	9	0.64
15	14	1	1	0.07
14	14	0	0	0
14	14	0	0	0
98			Total	2.14

Level of significance :

Let $\alpha = 0.05$

Critical value: The table value of χ^2 at $\alpha = 0.05$ is for $(7 - 1)(2 - 1)$ i.e.; 6 degree is $\chi^2 = 12.592$

Since the calculated value of χ^2 (7.16) < Table value of χ^2 at level of significance 0.05 for six degree = 12.592

Hence, the null Hypothesis is accepted i.e.; the air craft accident are uniformly distributed over the week. **Ans.**

62.30 CHI-SQUARE TEST AS A TEST OF INDEPENDENCE

$$\text{Expected Frequency} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

Example 16. *In an investigation into the health and nutrition of two groups of children of different social status the following results are obtained.*

Health \ *Social Status*	*Poor*	*Rich*	*Total*
Below Normal	*130*	*20*	*150*
Normal	*102*	*108*	*210*
Above normal	*24*	*96*	*120*
Total	*256*	*224*	*480*

Discuss the relation between the health and their social status.

Solution. H_0. Null Hypothesis : There is no association between health and social status.

H_1. Alternate Hypothesis: there an no association between health and social status.

$$\text{Expected frequency} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

Health \ *Social Status*	*Poor*	*Rich*	*Total*
Below Normal	$\frac{256 \times 150}{480} = 80$	$\frac{224 \times 150}{480} = 70$	150
Normal	$\frac{256 \times 210}{480} = 112$	$\frac{224 \times 210}{480} = 98$	210
Above normal	$\frac{256 \times 120}{480} = 64$	$\frac{224 \times 120}{480} = 56$	120
Total	256	224	480

Total number of observed frequencies = Total number of expected frequencies = 480

Degree of freedom = (3 − 1) (2 − 1) = 2

Level of significance, take $\alpha = 0.5$

Critical value the table value of χ^2 at $\alpha = 0.05$ for degree of freedom 2 is 5.99.

Decision. Since the calculated value of χ^2 (122.44) > table value of χ^2 at level of significance 0.05 for two $2 d.f.$ = 5.991.

Hence, the null hypothesis is rejected i.e., social status and health are associated (Dependent). **Ans.**

Calculation of Chi-square

Observed value (O)	Expected Value (E)	$O - E$	$(O - E)^2$	$\frac{(O-E)^2}{E}$
130	80	50	2500	$\frac{2500}{80} = 31.25$
102	112	– 10	100	$\frac{100}{112} = 0.89$
24	64	– 40	1600	$\frac{1600}{64} = 25$
20	70	– 50	2500	$\frac{2500}{70} = 35.71$
108	98	10	100	$\frac{100}{98} = 1.02$
96	56	40	1600	$\frac{1600}{56} = 28.57$
			Total	122.44

Example 17. *The I.Q. and economic condition of home of 1000 students of an engineering college, Delhi were noted as given in the table :*

I.Q. / *Economic con.*	*High*	*Low*	*Total*
Rich	*100*	*300*	*400*
Poor	*350*	*250*	*600*
Total	*450*	*550*	*1,000*

Find out whether there is any association between economic condition at home and I.Q. of the students.

Given for 1 *d.f.*, χ^2 *at the level of significance 0.05 is 3.84.*

Solution.

Null Hypothesis H_0: There is no association between economic condition at home and I.Q.

Alternative hypothesis H_1: There is an association between economic condition at home and I . Q.

$$\text{Expected frequency E } = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

I.Q. / *Economic cond.*	*High*	*Low*	*Total*
Rich	$\frac{400 \times 450}{1000} = 180$	$\frac{400 \times 550}{1000} = 220$	400
Poor	$\frac{600 \times 450}{1000} = 270$	$\frac{600 \times 550}{1000} = 330$	600
Total	450	550	1000

Calculation of Chi-square

Observed value (O)	Expected Value (E)	$O-E$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
100	180	– 80	6400	35.5
350	270	80	6400	23.7
300	220	80	6400	29.1
250	330	– 80	6400	19.4
			Total	107.7

Degree of freedom = (R – 1) (C – 1) = (2 – 1) (2 – 1) = 1 given for *d.f.* = 1,
χ^2 at the level of significance 0.05 = 3.84

Decision. The calculated value of χ^2 is greater than Table value of χ^2. Hence, the hypothesis is rejected and the alternative hypothesis is accepted.

Hence, there is an association between economic condition at home and I.Q. **Ans.**

Example 18. *To test the effectiveness of inoculation against cholera, the following table was obtained.*

	Attached	*Not attached*	*Total*
Inoculated	30	160	190
Notinoculated	140	460	600
Total	170	620	790

(*The figures represent the number of persons*)

Use χ^2 - test to defend or refute the statement. The inoculation prevents attack from cholera. (*U.P. III Semester Dec. 2009*)

Solution. **H_0 Null Hypothesis:** No inoculation prevents attack from cholera.

H_1 Alternate Hypothesis:

The inoculation prevents attack from cholera.

Expected frequency $E = \dfrac{\text{Row total} \times \text{Column Total}}{\text{Grand total}}$

	Attacked	Not Attacked	Total
Inoculated	$\frac{170 \times 190}{790} = 40.9$	$\frac{620 \times 190}{790} = 149.1$	190
Not inoculated	$\frac{170 \times 600}{790} = 129.1$	$\frac{620 \times 600}{790} = 470.9$	600
Total	170	620	790

Total number of observed frequencies

= Total number of expected frequencies = 790

Calculation of Chi-square

Observed value (O)	Expected value (E)	(O – E)	$(O - E)^2$	$\frac{(O-E)^2}{E}$
30	40.9	– 10.9	118.81	2.904
140	129.1	10.9	118.81	0.920
160	149.1	10.9	118.81	0.797
460	470.9	– 10.9	118.81	0.252
			Total	4.873

Degree of freedom = (R – 1) (C – 1) = (2 – 1) (2 – 1) = 1

The critical value of the table value of χ^2 at $\alpha = 0.05$ for 1d.f. is 3.841.

Decision: Since the calculated value of χ^2 (4.873) is greater than the table value (3.841). Thus, the hypothesis is rejected and the alternative hypothesis is accepted. Hence, the inoculation prevents attack from cholera. **Ans.**

EXERCISE 62.2

1. In an experiment immunization of cattle from a disease, the following results are obtain:

	Affected	Unaffected	Total
Inoculated	12	28	40
No Inoculated	13	7	20
Total	25	35	60

Examine the effect of vaccin in controlling the incidence of the disease. **Ans.** Not independent

2. In the contigency table given below use Chi-square test to test for independence of hair colour and eye colour of persons:

Hair colour / Eye colour	Light	Dark	Total
Blue	26	9	35
Brown	7	18	25
Total	33	27	60

Ans. Hair colour and eye colour are associated

3. A survey amongst women was conducted to study the family life. The observations are as follows:

Family life

	Happy	Not Happy	Total
Educated	70	30	100
Not educated	60	40	100
Total	130	70	200

Test whether there is any association between family life and education.

Ans. there is no association between family life and education.

4. A certain drug was administrated to 500 people out of a total of 800 included in a sample to test its efficiency against typhoid, the results are given below :

	Typhoid	No Typhoid	Total
Drug	200	300	500
No Drug	280	20	300
	48	320	800

On the basis of the data, can we say that drug is effective in preventing Typhoid.

Ans. $\chi^2 = 222.22$ drug is effective

5. The following table gives the number of person's whose eye sight is attacked and an injection of macugen is injected by Prof. Atul.

	Eye sight Improved	Eye sight not improved	total
Injected	216	145	361
Not injected	105	234	339
Total	321	379	700

Do you think macugen injection can improve the eye sight. **Ans.** By injection, eye sight has improved

6. From the table given below, whether the colour of the sons eyes is associated with that of father's eye.

Eyes colour in sons

		Not light	light	
Eyes colour in fathers	Not light	230	148	378
	light	151	471	622
		381	619	1000

There is an association between the colour of eyes of sons and colours of eye's of fathers.

Ans. Null hypothesis in rejected

7. The following table gives the classification of 500 plants according to the nature of leaf and flower colour.

	Blue flowers	White flower	Total
Flat leaf	329	121	450
Creppled leaf	78	32	110
Total	407	153	560

Test whether they have any association between them. **Ans.** No association between them

8. The table below gives the data obtained from a hospital of sugar patients:

	Cured	Not cured	Total
Inoculated	31	469	500
Not inoculated	185	1,315	1500
	216	1784	2000

Test the effectiveness of inoculation in preventing the sugar disease.

Ans. Inoculation is effective in preventing the attack of sugar

9. The following table give the no. of good and defective parts produced by each of these shifts in a factory. Test whether the shift has any association with good or defective parts.

	Day shift	Evening shift	Night shift	Total
Good parts	960	940	950	2850
Defective parts	40	50	45	135
	1000	990	995	2985

Ans. χ^2 = 0.3485, Null Hypothesis is accepted

10. The following table shows the results of drug against B.P.

	Not attacked	Attacked	Total
Drug	267	37	304
No Drug	757	155	912
Total	1024	192	1216

Find out whether there is any significance association between drug and attack.

Ans. Drug prevents the attack of B.P.

CHAPTER

63

BINOMIAL DISTRIBUTION

63.1 BINOMIAL DISTRIBUTION P (r) = $^nC_r\ p^r \cdot q^{n-r}$

To find the probability of the happening of an event once, twice, thrice, ... r times ...... exactly in n trials.

Let the probability of the happening of an event A in one trial be p and its probability of not happening be $1 - p = q$.

We assume that there are n trials and the happening of the event A is r times and its not happening is $n - r$ times.

$$\underbrace{A\ A ....A}_{r \text{ times}} \qquad \underbrace{\bar{A}.\bar{A}......\bar{A}}_{n-r \text{ times}} \qquad ...(1)$$

A indicates its happening, $\bar{A}$ its failure and $P(A) = p$ and $P(\bar{A}) = q$.

We see that (1) has the probability

$$\underbrace{pp\ ....\ p}_{r \text{ times}}\ \underbrace{q\ .\ q\ ....\ q}_{n-r \text{ times}} = p^r \cdot q^{n-r} = p^r q^{n-r} \qquad ...(2)$$

Clearly (1) is merely one order of arranging $r\,A$'s.

The probability of (1) $= p^r q^{n-r} \times$ Number of different arrangements of $r\,A$'s and $(n-r)\,\bar{A}$'s.

The number of different arrangements of $r\,A$'s and $(n-r)\,\bar{A}$'s $= {}^nC_r$

$\therefore$ Probability of the happening of an event r times ${}^nC_r\ p^r \cdot q^{n-r}$.

$$P(r) = {}^nC_r\ p^r \cdot q^{n-r} \quad (r = 0, 1, 2, ...,n).$$
$$= (r+1)\text{th term of } (q+p)^n$$

If $r = 0$, probability of happening of an event 0 times $= {}^nC_0\ q^n p^0 = q^n$

If $r = 1$, probability of happening of an event 1 time $= {}^nC_1\, q^{n-1}\, p$

If $r = 2$, probability of happening of an event 2 times $= {}^nC_2\ q^{n-2}\, p^2$

If $r = 3$, probability of happening of an event 3 times $= {}^nC_3\ q^{n-3}\, p^3$ and so on.

These terms are clearly the successive terms in the expansion of $(q+p)^n$.

Hence it is called Binomial Distribution.

Example 1. *Find the probability of getting 4 heads in 6 tosses of a fair coin.*

Solution. $p = \frac{1}{2}$, $q = \frac{1}{2}$, $n = 6, r = 4$.

We know that $\quad P(r) = {}^nC_r\ q^{n-r}\ p^r$

$$P(4) = {}^6C_4\ q^{6-4}\ p^4 = \frac{6\times5}{1\times2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4 = 15\times\left(\frac{1}{2}\right)^6 = \frac{15}{64}$$ **Ans.**