CHAPTER THREE

Theory of Probability-I

LEARNING OBJECTIVES. Upon completion of this chapter, you should be able to :

1. **Learn about the different approaches to the theory of probability like the classical/ mathematical probability, empirical probability and axiomatic probability, along with their limitations.**
2. **Understand the addition theorem, the multiplication theorem and other important theorems on probability , and their use in solving problems in various diversified situations.**
3. **Demonstrate the concepts of conditional probability, pairwise independence and mutual independence of events**
4. **Appreciate the use of probability theory in our day-to-day life and in the decision-making in the face of uncertainty.**

CHAPTER OUTLINE

3·1. INTRODUCTION
3·2. SHORT HISTORY
3·3. BASIC TERMINOLOGY
3·4. MATHEMATICAL (OR CLASSICAL OR 'A PRIORI') PROBABILITY
  3·4·1. Limitations of Mathematical Probability
3·5. STATISTICAL (OR EMPIRICAL) PROBABILITY
  3·5·1. Limitations of Empirical Probability
3·6. SUBJECTIVE PROBABILITY
3·7. MATHEMATICAL TOOLS : PRELIMINARY NOTIONS OF SETS
  3·7·1. Sets and Elements of Sets
  3·7·2. Operations on Sets
  3·7·3. Algebra of Sets
  3·7·4. Limit of Sequence of Sets
  3·7·5. Classes of Sets.
3·8. AXIOMATIC APPROACH TO PROBABILITY
  3·8·1. Random Experiment, Sample Space and Elementary Events
  3·8·2. Event
  3·8·3. Acceptable Assignment of Probabilities
  3·8·4. Natural Assignment of Probabilities

correspondence between themselves ultimately solved this problem and this correspondence laid the first foundation of the science of probability. Next stalwart in this field was J. Bernoulli (1654-1705) whose *'Treatise on Probability'* was published posthumously by his nephew N. Bernoulli in 1713. De-Moivre (1667-1754) also did considerable work in this field and published his famous *'Doctrine of Chances'* in 1718. Other main contributors are : T. Bayes (Inverse probability), P.S. Laplace (1749-1827) who after extensive research over a number of years finally published *'Theoric analytique des probabilities'* in 1812. In addition to these, other outstanding contributors are Levy, Mises and R.A. Fisher.

Russian mathematician also have made very valuable contributions to the modern theory of probability. Chief contributors, to mention only a few of them are : Chebychev (1821-94) who founded the Russian School of Statisticians; A. Markoff (1856-1922); Liapounoff (Central Limit Theorem) ; A. Khintchine (Law of Large Numbers) and A. Kolmogorov, who axiomised the calculus of probability.

3.3. BASIC TERMINOLOGY

In this section we shall explain the various terms which are used in the definition of probability under different approaches.

1. *Random Experiment.* *If in each trial of an experiment conducted under identical conditions, the outcome is not unique, but may be any one of the possible outcomes, then such an experiment is called a random experiment.*

Examples of random experiments are : tossing a coin, throwing a die, selecting a card from a pack of playing cards, selecting a family out of a given group of families, etc. In all these cases, there are a number of possible results which can occur but there is an uncertainty as to which one of them will actually occur.

Notes : (*i*) A *die* is a small cube used in gambling. On its six faces, dots are marked as

Plural of die is *dice*. The outcome of throwing a die is the number of dots on its uppermost face.

(*ii*) A *pack of cards* consists of four suits called *Spades, Hearts, Diamonds* and *Clubs*. Each suit consists of 13 cards, of which nine cards are numbered from 2 to 10, an ace, king, a queen and a jack (or knave). Spades and clubs are black-faced cards, while hearts and diamonds are red-faced cards.

2. *Outcome.* The result of a random experiment will be called an *outcome*.

3. *Trial and Event.* Any particular performance of a random experiment is called a *trial* and outcome or combination of outcomes are termed as *events*. For example,

(*i*) If a coin is tossed repeatedly, the result is not unique. We may get any of the two faces, head or tail. Thus tossing of a coin is a random experiment or trial and getting of a head or tail is an event.

(*ii*) In an experiment which consists of the throw of a six-faced die and observing the number of points that appear, the possible outcomes are 1, 2, 3, 4, 5, 6

In the same experiment, the possible events could also be stated as

'Odd number of points' ; *'Even number of points'*; *'Getting a point greater than 4'* ; and so on.

Event is called *simple* if it corresponds to a single possible outcome of the experiment otherwise it is known as a *compound* or *composite* event. Thus in tossing of a single die the event of getting '6' is a simple event but the event of getting an even number is a composite event.

4. *Exhaustive Events or Cases.* The total number of possible outcomes of a random experiment is known as the *exhaustive events or cases*. For example,

(*i*) In tossing of a coin, there are two exhaustive cases, *viz.*, head and tail (the possibility of the coin standing on an edge being ignored).

(*ii*) In throwing of a die, there are 6 exhaustive cases since any one of the 6 faces 1, 2, ..., 6 may come uppermost.

(*iii*) In drawing two cards from a pack of cards, the exhaustive number of cases is ${}^{52}C_2$, since 2 cards can be drawn out of 52 cards in ${}^{52}C_2$ ways.

(*iv*) In throwing of two dice, the exhaustive number of cases is $6^2 = 36$, since any of the numbers 1 to 6 on the first die can be associated with any of the 6 numbers on the other die. In general, in throwing of n dice, the exhaustive number of cases is 6^n.

5. *Favourable Events or Cases.* The number of cases favourable to an event in a trial is the number of outcomes which entail the happening of the event. For example,

(*i*) In drawing a card from a pack of cards the number of cases favourable to drawing of an ace is 4, for drawing a spade is 13 and for drawing a red card is 26.

(*ii*) In throwing of two dice, the number of cases favourable to getting the sum 5 is :

(1, 4), (4, 1), (2, 3), (3, 2), *i.e.*, 4.

6. *Mutually Exclusive Events.* Events are said to be *mutually exclusive* or *incompatible* if the happening of any one of them precludes the happening of all the others, *i.e.*, if no two or more of them can happen simultaneously in the same trial. For example,

(*i*) In throwing a die all the 6 faces numbered 1 to 6 are mutually exclusive since if any one of these faces comes, the possibility of others, in the same trial, is ruled out.

(*ii*) Similarly in tossing a coin the events head and tail are mutually exclusive.

7. *Equally Likely Events.* Outcomes of trial are said to be *equally likely* if taking into consideration all the relevant evidences, there is no reason to expect one in preference to the others. For example,

(*i*) In a random toss of an unbiased or uniform coin, head and tail are equally likely events.

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

8. *Independent Events.* Several events are said to be independent if the happening (or non-happening) of an event is not affected by the supplementary knowledge concerning the occurrence of any number of the remaining events. For example,

(*i*) In tossing an unbiased coin, the event of getting a head in the first toss is independent of getting a head in the second, third and subsequent throws.

(*ii*) When a die is thrown twice, the result of the first throw does not affect the result of the second throw.

(*iii*) If we draw a card from a pack of well-shuffled cards and replace it before drawing the second card, the result of the second draw is independent of the first draw. But, however, if the first card drawn is not replaced then the second draw is dependent on the first draw.

3·4. MATHEMATICAL (OR CLASSICAL OR 'A PRIORI') PROBABILITY

Definition. *If a random experiment or a trial results in 'n' exhaustive, mutually exclusive and equally likely outcomes (or cases), out of which m are favourable to the occurrence of an event E, then the probability 'p' of occurrence (or happening) of E, usually denoted by P(E), is given by :*

$$p = P(E) = \frac{\text{Number of favourable cases}}{\text{Total number of exhaustive cases}} = \frac{m}{n} \qquad \ldots (3\text{·}1)$$

This definition was given by James Bernoulli who was the first person to obtain a quantitative measure of uncertainty.

Remarks **1.** Since $m \geq 0$, $n > 0$ and $m \leq n$, we get from (3·1) :

$$P(E) \geq 0 \quad \text{and} \quad P(E) \leq 1 \quad \Rightarrow \quad 0 \leq P(E) \leq 1$$

2. Sometimes we express (3·1) by saying that 'the odds in favour of E are $m : (n - m)$ or the odds against E are $(n - m) : m$.

3. The non-happening of the event E is called the *complementary event of* E and is denoted by $\bar{E}$ or E^c. The number of cases favourable to $\bar{E}$, *i.e.*, non-happening of E is $(n - m)$. Then the probability q that E will not happen is given by :

$$q = P(\bar{E}) = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - p \quad \Rightarrow \quad p + q = 1 \qquad \ldots (3\text{·}1a)$$

$$\therefore \quad q = P(\bar{E}) = 1 - P(E) \quad \Rightarrow \quad P(E) = 1 - P(\bar{E}) \quad \text{or} \quad P(E) + P(\bar{E}) = 1 \qquad \ldots (3\text{·}1b)$$

If the event E represents the happening of at least one of the events $E_1, E_2, \ldots, E_n$, then its complementary event $\bar{E}$ represents the happening of none of the events $E_1, E_2, \ldots, E_n$. Substituting in (3·1 *b*), we get

$$P(\text{Happening of at least one of the events } E_1, E_2, \ldots, E_n)$$
$$= 1 - P(\text{None of the events } E_1, E_2, \ldots, E_n \text{ happens}) \qquad \ldots (3\text{·}1c)$$

4. Probability 'p' of the happening of an event is also known as the probability of success and the probability 'q' of the non-happening of the event as the probability of failure, *i.e.*, $(p + q = 1)$.

5. If $P(E) = 1$, E is called a *certain event* and if $P(E) = 0$, E is called an *impossible* event.

6. We can compute the probability in (3·1) by logical reasoning, without conducting any experiment. Since, the probability in (3·1) can be computed prior to obtaining any experimental data, it is also termed as *'a priori'* or *mathematical probability*.

3·4·1. Limitations of Classical Definition. This definition of classical probability breaks down in the following cases :

(*i*) If the various outcomes of the random experiment are not equally likely or equally probable. For example,

(*a*) The probability that a candidate will pass in a certain test is not 50%, since the two possible outcomes, *viz.*, success and failure (excluding the possibility of a compartment) are not equally likely.

(*b*) The probability that a ceiling fan in a room will fall is not 1/2, since the events of the fan 'falling' and 'not falling' though mutually exclusive and exhaustive, are not equally likely. In fact, the probability of the fan falling will be almost zero.

(*c*) If a person jumps from a running train, then the probability of his survival will not be 50%, since in this case the events survival and death, though exhaustive and mutually exclusive, are not equally likely.

(*ii*) If the exhaustive number of outcomes of the random experiment is infinite or unknown.

3·5. STATISTICAL (OR EMPIRICAL) PROBABILITY

Definition. *(VON MISES). If an experiment is performed repeatedly under essentially homogeneous and identical conditions, then the limiting value of the ratio of the number of times the event occurs to the number of trials, as the number of trials becomes indefinitely large, is called the probability of happening of the event, it being assumed that the limit is finite and unique.*

Symbolically, if in N trials an event E happens M times, then the probability of the happening of E, denoted by $P(E)$, is given by :

$$P(E) = \lim_{N \to \infty} \frac{M}{N} \qquad \ldots (3\cdot2)$$

Remarks **1.** Since in the relative frequency approach, the probability is obtained objectively by repetitive empirical observations, it is also known as *'Empirical Probability'*.

2. An experiment is unique and non-repeating only in the case of *subjective probability*. In other cases, there are a large number of experiments or trials to establish the chance of occurrence of an event. This is particularly so in case of empirical probability. In classical probability also, repeated experiments may be made to verify whether a deduction on the basis of certain axioms or undisputed laws is justified. Only after repeated trials it can be established that the chance of head in a toss of a coin is 1/2. J. E. Kerrich conducted coin tossing experiment with 10 sets of 1,000 tosses each during his confinement in World War II. The number of heads found by him were : 502, 511, 497, 529, 504, 476, 507, 520, 504, 529.

This gives the probability of getting a head in a toss of a coin as : $\frac{5,079}{10,000} = 0{\cdot}5079 \simeq \frac{1}{2}$.

Thus, the empirical probability approaches the classical probability as the number of trials becomes indefinitely large.

3·5·1. Limitations of Empirical Probability. (*i*) If an experiment is repeated a large number of times, the experimental conditions may not remain identical and homogeneous.

(*ii*) The limit in (3·2) may not attain a unique value, however large N may be.

Example 3·1. *What is the chance that a leap year selected at random will contain 53 Sundays ?*

Solution. In a leap year (which consists of 366 days), there are 52 complete weeks and 2 days over. The following are the possible combinations for these two *'over'* days : (*i*) Sunday and Monday, (*ii*) Monday and Tuesday, (*iii*) Tuesday and Wednesday, (*iv*) Wednesday and Thursday, (*v*) Thursday and Friday, (*vi*) Friday and Saturday, and (*vii*) Saturday and Sunday.

In order that a leap year selected at random should contain 53 Sundays, one of the two 'over' days must be Sunday. Since out of the above 7 possibilities, 2, *viz.*, (*i*) and (*vii*), are favourable to this event.

$\therefore$ Required probability $= \frac{2}{7}$.

Example 3·2. *Two unbiased dice are thrown. Find the probability that :*

(i) both the dice show the same number,

(ii) the first die shows 6,

(iii) the total of the numbers on the dice is 8,

(iv) the total of the numbers on the dice is greater than 8,

(v) the total of the numbers on the dice is 13, and

(vi) the total of the numbers on the dice is any number from 2 to 12, both inclusive.

Solution. In a random throw of two dice, since each of the six faces of one die can be associated with each of six faces of the other die, the total number of cases is $6 \times 6 = 36$, as given below :

(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)
(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

Here, the expression, say, (i, j) means that the first die shows the number i and the second die shows the number j. Obviously, $(i, j) \neq (j, i)$ if $i \neq j$.

$\therefore$ Exhaustive number of cases $(n) = 36$.

(a) The favourable cases that both the dice show the same number are :

(1, 1), (2, 2), (3, 3), (4, 4), (5, 5) and (6, 6), *i.e.*, $m = 6$.

$\therefore$ Probability that the two dice show the same number $= \frac{6}{36} = \frac{1}{6}$.

(b) The favourable cases that the first die shows 6 are :

(6, 1), (6, 2), (6, 3), (6, 4), (6, 5) and (6, 6), *i.e.*, 6 in all.

$\therefore$ Probability that the first die shows '6' $= \frac{6}{36} = \frac{1}{6}$.

(c) The cases favourable to getting a total of 8 on the two dice are :

(2, 6), (3, 5), (4, 4), (5, 3), (6, 2), *i.e.*, $m = 5$.

$\therefore$ Probability that total of numbers on two dice is $8 = \frac{5}{36}$.

(d) The cases favourable to getting a total of more than 8 are :

(3, 6), (6, 3), (4, 5), (5, 4), (4, 6), (6, 4), (5, 5), (5, 6), (6, 5), (6, 6), *i.e.*, $m = 10$.

$\therefore$ Probability that the total of numbers on two dice is greater than $8 = \frac{10}{36} = \frac{5}{18}$.

(e) This is an example of an *impossible event*, since the maximum total can be $6 + 6 = 12$. Therefore, the required probability is 0.

(f) The probability is 1, as the total of the numbers on the two dice certainly ranges from 2 to 12. The given event is called a *certain event*.

Example 3·3. *(a) Among the digits 1, 2, 3, 4, 5 at first one is chosen and then a second selection is made among the remaining four digits. Assuming that all twenty possible outcomes have equal probabilities, find the probability that an odd digit will be selected*

(i) the first time, (ii) the second time, and (iii) both times.

(b) From 25 tickets, marked with first 25 numerals, one is drawn at random. Find the chance that (i) it is multiple of 5 or 7, and (ii) it is a multiple of 3 or 7.

Solution. *(a)* Total number of cases $= 5 \times 4 = 20$.

(i) Now there are 12 cas
(1, 4), (1, 5), (3, 1), (3, 2), (3,

$\therefore$ The probability tha

(ii) Also there are 12 c
(3, 1), (4, 1), (5, 1), (1, 3), (2, 3

$\therefore$ The probability tha

(iii) There are six case
(1, 5), (3, 1), (3, 5), (5, 1) and

$\therefore$ The probability tha

(b) (i) Numbers (out of
20 and 25, *i.e.*, 5 in all and t
in all. Hence required numb

$\therefore$ Required probab

(ii) Numbers (among th
15, 18, 21, 24, *i.e.*, 8 in all ; an
in all. Since the number 2
distinct favourable cases is

$\therefore$ Require

Example 3·4. *(a) Four*
probability that

(i) They are a king, a

(ii) Two are kings and

(iii) Two are black and

(iv) There are two card

(b) In shuffling a pack of
missing cards should be one fr

Solution. Four cards c
ways, which gives the exha

(i) 1 king can be drawn
and an ace can each be dra
king can be associated with
the favourable number of c

Hence the required pr

(ii) Required pr

(iii) Since there are 26
diamonds and hearts) in a p

(iv) Required pr

(*i*) Now there are 12 cases in which the first digit drawn is odd, *viz.*, (1, 2), (1, 3), (1, 4), (1, 5), (3, 1), (3, 2), (3, 4), (3, 5), (5, 1), (5, 2), (5, 3) and (5, 4).

$\therefore$ The probability that the first digit drawn is odd $= \frac{12}{20} = \frac{3}{5}$.

(*ii*) Also there are 12 cases in which the second digit drawn is odd, *viz.*, (2, 1), (3, 1), (4, 1), (5, 1), (1, 3), (2, 3), (4, 3), (5, 3), (1, 5), (2, 5), (3, 5) and (4, 5).

$\therefore$ The probability that the second digit drawn is odd $= \frac{12}{20} = \frac{3}{5}$.

(*iii*) There are six cases in which both the digits drawn are odd, *viz.*, (1, 3), (1, 5), (3, 1), (3, 5), (5, 1) and (5, 3).

$\therefore$ The probability that both the digits drawn are odd $= \frac{6}{20} = \frac{3}{10}$.

(*b*) (*i*) Numbers (out of the first 25 numerals) which are multiples of 5 are 5, 10, 15, 20 and 25, *i.e.*, 5 in all and the numbers which are multiples of 7 are 7, 14 and 21, *i.e.*, 3 in all. Hence required number of favourable cases are 5 + 3 = 8.

$\therefore$ Required probability $= \frac{8}{25}$.

(*ii*) Numbers (among the first 25 numerals) which are multiples of 3 are 3, 6, 9, 12, 15, 18, 21, 24, *i.e.*, 8 in all ; and the numbers, which are multiples of 7 are 7, 14, 21, *i.e.*, 3 in all. Since the number 21 is common in both the cases, the required number of distinct favourable cases is $8 + 3 - 1 = 10$.

$\therefore$ Required probability $= \frac{10}{25} = \frac{2}{5}$.

Example 3·4. *(a) Four cards are drawn at random from a pack of 52 cards. Find the probability that*

(i) They are a king, a queen, a jack and an ace.

(ii) Two are kings and two are queens.

(iii) Two are black and two are red.

(iv) There are two cards of hearts and two cards of diamonds.

(b) In shuffling a pack of cards, four are accidentally dropped, find the chance that the missing cards should be one from each suit.

Solution. Four cards can be drawn from a well-shuffled pack of 52 cards in ${}^{52}C_4$ ways, which gives the exhaustive number of cases.

(*i*) 1 king can be drawn out of the 4 kings in ${}^4C_1$ ways. Similarly, 1 queen, 1 jack and an ace can each be drawn in ${}^4C_1 = 4$ ways. Since any one of the ways of drawing a king can be associated with any one of the ways of drawing a queen, a jack and an ace, the favourable number of cases are ${}^4C_1 \times {}^4C_1 \times {}^4C_1 \times {}^4C_1$.

Hence the required probability $= \frac{{}^4C_1 \times {}^4C_1 \times {}^4C_1 \times {}^4C_1}{{}^{52}C_4} = \frac{256}{{}^{52}C_4}$.

(*ii*) Required probability $= \frac{{}^4C_2 \times {}^4C_2}{{}^{52}C_4}$.

(*iii*) Since there are 26 black cards (of spades and clubs) and 26 red cards (of diamonds and hearts) in a pack of cards, the required probability $= \frac{{}^{26}C_2 \times {}^{26}C_2}{{}^{52}C_4}$.

(*iv*) Required probability $= \frac{{}^{13}C_2 \times {}^{13}C_2}{{}^{52}C_4}$.

In all the above cases, the probabilities (chances) of occurrence of the corresponding events are assigned by individuals and are based on their personal judgement, wisdom, intuition and expertise. These probabilities are called the *subjective probabilities* and represent the degree of belief and the confidence, one has in the occurrence of the respective event. Such probabilities can be calculated in all the cases, though the probabilities assigned by different persons to the same event may vary from person to person.

3.7. MATHEMATICAL TOOLS : PRELIMINARY NOTIONS OF SETS

The set theory was developed by the German mathematician, G. Cantor (1845-1918).

3·7·1. Sets and Elements of Sets. A set is a well-defined collection or aggregate of all possible objects having given properties and specified according to a well-defined rule. The objects comprising a set are called elements, members or points of the set. Sets are often denoted by capital letters, *viz.*, A, B, C, etc. If x is an element of the set A, we write symbolically $x \in A$ (x belongs to A). If x is not a member of the set A, we write $x \notin A$ (x does not belong to A). Sets are often described by describing the properties possessed by their members. Thus the set A of all non-negative rational numbers with square less than 2 will be written as :

$$A = \{x : x \text{ rational}, x \geq 0, x^2 < 2\}.$$

If every element of the set A belongs to the set B, *i.e.*, if $x \in A \Rightarrow x \in B$ then we say that A is a subset of B and write symbolically $A \subseteq B$ (A is contained in B) or $B \supseteq A$ (B contains A). Two sets A and B are said to be *equal* or *identical* if $A \subseteq B$ and $B \subseteq A$ and write $A = B$ or $B = A$.

A *null* or an *empty* set is one which does not contain any element at all and is denoted by ϕ.

Remarks. 1. Every set is a subset of itself.

2. An empty set is subset of every set.

3. A set containing only one element is conceptually distinct from the element itself, but will be represented by the same symbol for the sake of convenience.

4. As will be the case in all our applications of set theory, especially to probability theory, we shall have a fixed set S, (say), given in advance, and we shall be concerned only with subsets of this given set. The underlying set S may vary from one application to another, it will be referred to as *universal set* of each particular discourse.

3·7·2. Operations on Sets. The *union* of two given sets A and B, denoted by $A \cup B$, is defined as a set consisting of all those points which belong to either A or B or both. Thus symbolically : $A \cup B = \{x : x \in A \text{ or } x \in B\}$.

Similarly $\bigcup_{i=1}^{n} A_i = \{x : x \in A_i \text{ for at least one } i = 1, 2, \ldots, n\}$

The *intersection* of two sets A and B, denoted by $A \cap B$, is defined as a set consisting of those elements which belong to both A and B. Thus

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

Similarly $\bigcap_{i=1}^{n} A_i = \{x : x \in A_i \text{ for all } i = 1, 2, \ldots, n\}$.

For example, if $A = \{1, 2, 5, 8, 10\}$ and $B = \{2, 4, 8, 12\}$, then

$A \cup B = \{1, 2, 4, 5, 8, 10, 12\}$ and $A \cap B = \{2, 8\}$.

If A and B have no common point, *i.e.*, $A \cap B = \phi$, then the sets A and B are said to be *disjoint, mutually exclusive* or *non-overlapping.*

The *relative difference* of a set A from another set B, denoted by $A - B$ is defined as a set consisting of those elements of A which do not belong to B. Symbolically

$$A - B = \{x : x \in A \text{ and } x \notin B\}.$$

The *complement* or *negative* of any set A, denoted by $\overline{A}$ is a set containing all elements of the universal set S (say) that are not elements of A, *i.e.*, $\overline{A} = S - A$.

3·7·3. Algebra of Sets. Now we state certain important properties concerning operations on sets. If A, B and C are the subsets of a universal set S, then the following laws hold :

Commutative Laws :
$$A \cup B = B \cup A, \qquad A \cap B = B \cap A$$

Associative Law :
$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

Distributive Law :
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Complementary Law :
$$A \cup \overline{A} = S, \quad A \cap \overline{A} = \phi$$
$$A \cup S = S, \quad \text{and} \quad A \cap S = A \quad (\because A \subset S),$$
$$A \cup \phi = A, \quad A \cap \phi = \phi.$$

Difference Law :
$$A - B = A \cap \overline{B}.$$
$$A - B = A - (A \cap B) = (A \cup B) - B$$
$$A - (B - C) = (A - B) \cup (A - C)$$
$$(A \cup B) - C = (A - C) \cup (B - C)$$
$$A - (B \cup C) = (A - B) \cap (A - C)$$
$$(A \cap B) \cup (A - B) = A, \quad (A \cap B) \cap (A - B) = \phi$$

De-Morgan's Law

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

More generally

$$\overline{\left(\bigcup_{i=1}^{n} A_i\right)} = \bigcap_{i=1}^{n} \overline{A_i} \quad \text{and} \quad \overline{\left(\bigcap_{i=1}^{n} A_i\right)} = \bigcup_{i=1}^{n} \overline{A_i}$$

Involution Law :
$$\overline{(\overline{A})} = A$$

Idempotency Law :
$$A \cup A = A, \qquad A \cap A = A$$

3·7·4. Limit of Sequence of Sets. Let $\{A_n\}$ be a sequence of sets in S. The *limit supremum* or *limit superior* of the sequence, usually written as lim sup A_n, is the set of all those elements which belong to A_n for infinitely many n. Thus

$$\lim_{n \to \infty} \sup A_n = \{x : x \in A_n \text{ for infinitely many } n\} \qquad \ldots (3\cdot3)$$

The set of all those elements which belong to A_n for all but a finite number of n is called *limit infinimum or limit inferior* of the sequence and is denoted by lim inf A_n.

Thus $\lim_{n\to\infty} \inf A_n = \{x : x \in A_n \text{ for all but a finite number of } n\}$ ... (3·3a)

The sequence $\{A_n\}$ is said to have a limit if and only if lim sup A_n = lim inf A_n and this common value gives the limit of the sequence.

Theorem 3·1. *lim sup* $A_n = \bigcap_{m=1}^{\infty} \left(\bigcup_{n=m}^{\infty} A_n \right)$

and lim inf $A_n = \bigcup_{m=1}^{\infty} \left(\bigcap_{n=m}^{\infty} A_n \right)$

Def. $\{A_n\}$ is a monotone (infinite) sequence of sets if

either $(i)\ A_n \subset A_{n+1} \quad \forall \quad n$ or $(ii)\ A_n \supset A_{n+1} \ \forall\, n.$

In the former case the sequence $\{A_n\}$ is said to be *non-decreasing (or increasing) sequence* and is usually expressed as $A_n \uparrow$ and in the latter case it is said to be *non-increasing (or decreasing) sequence* and is expressed as $A_n \downarrow$.

For a monotone sequence (non-increasing or non-decreasing), the limit always exists and we have :

$$\lim_{n\to\infty} A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n \text{ in case } (i), \text{ i.e., } A_n \uparrow \\ \bigcap_{n=1}^{\infty} A_n \text{ in case } (ii), \text{ i.e., } A_n \downarrow \end{cases}$$

3·7·5. Classes of Sets. A group of sets will be termed as a *class* (of sets). Below we shall define some useful types of classes.

A ring **R** is a *non-empty* class of sets which is closed under the formation of finite 'unions' and 'difference',

i.e., if $A \in \mathbf{R}$, $B \in \mathbf{R}$, then $A \cup B \in \mathbf{R}$ and $A - B \in \mathbf{R}$.

Obviously ϕ is a member of every ring.

A *field* **F** (or an *algebra*) is a non-empty class of sets which is closed under the formation of finite unions and under complementation. Thus

$$(i)\ A \in \mathbf{F}, B \in \mathbf{F} \Rightarrow A \cup B \in \mathbf{F} \quad \text{and} \quad (ii)\ A \in \mathbf{F} \Rightarrow \bar{A} \in \mathbf{F}.$$

A σ-ring **C** is a non-empty class of sets which is closed under the formation of 'countable unions' and 'difference'. Thus

$$(i)\ A_i \in \mathbf{C}, i = 1, 2, \ldots \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathbf{C} \quad (ii)\ A \in \mathbf{C}, B \in \mathbf{C} \Rightarrow A - B \in \mathbf{C}.$$

More precisely σ-ring is a ring which is closed under the formation of countable unions.

A σ-field (or σ-algebra) **B** is a non-empty class of sets that is closed under the formation of 'countable unions' and complementations, *i.e.*,

$$(i)\ A_i \in \mathbf{B}, i = 1, 2, \ldots \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathbf{B}. \quad (ii)\ A \in \mathbf{B} \Rightarrow \bar{A} \in \mathbf{B}.$$

σ-field may also be defined as a field which is closed under the formation of countable unions.

Since, $P(S) = \sum_{i=1}^{N} P(e_i) = 1$, on using (3·4), we get

$$P(e_i) = \frac{1}{N};\ i = 1, 2, \ldots, N \qquad \ldots (3\cdot5)$$

in the case of natural assignment of probabilities.

For example, in the case of random toss of a coin, $S = (H, T)$ and the natural assignment of probabilities gives : $P(H) = P(T) = \frac{1}{2}$.

Similarly, in a random throw of a die, $S = \{1, 2, \ldots, 6\} = \{e_1, e_2, \ldots, e_6\}$, (say) and the natural assignment of probabilities is : $P(e_i) = \frac{1}{6};\ i = 1, 2, \ldots, 6$.

In the problems relating to random throw(s) of coin(s) or die (dice), the natural assignment of probabilities is possible only if the coins/dice are fair, uniform or unbiased. Similarly, in drawing cards from a well-shuffled pack of cards, for natural assignment of probabilities, all the draws must be made at random.

In the following discussion, whenever we use the adjective such as : *'fair coin'*, *'fair* or *unbiased die'*, card or a number drawn at *'random'*, etc., we will assume that the natural assignment is understood. Also if no assignment of probabilities is given in any question, the natural assignment is to be taken.

3·8·5. Axiomatic Probability. A purely mathematical definition of probability cannot give us the actual value of $P(A)$, the probability of occurrence of the event A and this must be considered as a function defined on all events. Accordingly, as in the case of any function we need a *domain space* which is the σ-field **B** of the events generated by S ; a *range space* which is the closed interval [0, 1] on the real line, and a rule which assigns a value to every element of the domain space **B**.

Definition. *Probability Function.* *$P(A)$ is the probability function defined on a σ-field* **B** *of events if the following properties or axioms hold.*

1. *For each $A \in$ **B**, $P(A)$ is defined, is real and $P(A) \geq 0$.* *(Axiom of non-negativity)*
2. $P(S) = 1$ *(Axiom of certainty)*
3. *If $\{A_n\}$ is any finite or infinite sequence of disjoint events in* **B**, *then*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \qquad \text{(Axiom of additivity)} \qquad (3\cdot6)$$

An obvious advantage of this definition is that it is not concerned with assigning numerical values to the probabilities of events.

Remarks **1.** The set function **P** defined on σ-field **B**, taking its values in the real line and satisfying the above three axioms is called the *probability measure.*

2. The same definition of probability applies to *uncountable sample space* except that special restrictions must be placed on S and its subsets. It is important to realise that for a complete description of a probability measure, three things must be specified, *viz.*, the sample space S, the σ-field (σ-algebra) **B** formed from certain subsets of S and set function **P**. The triplet $(S, \mathbf{B}, \mathbf{P})$ is often called the *probability space.* In most elementary applications, S is finite and the σ-algebra **B** is taken to be the collection of all subsets of S.

3. ***Classical Probability as a Special Case of Axiomatic Probability.*** Let us consider a random experiment whose sample space S has finite number of points, say N, given by :

$$S = \{e_1, e_2, \ldots, e_N\},$$

where obviously $e_1, e_2, \ldots, e_N$ are mutually disjoint events.

Let us now suppose that all the outcomes of the random experiment are equally likely so that the probability of each of the singleton events $\{e_i\}$ is $1/N$, *i.e.*,

$$P\{e_i\} = \frac{1}{N};\ i = 1, 2, \ldots, n \qquad \ldots (*)$$

Further, let an event A consist of k distinct point of S, say $e_{i1}, e_{i2}, \ldots, e_{ik}$. Then by the axiom of additivity (Axiom 3), we get

$$P(A) = \sum_{j=1}^{k} P\{e_{ij}\} = \sum_{j=1}^{k} \frac{1}{N} = \frac{k}{N} \qquad \text{[Using (*)]}$$

$$= \frac{\text{Number of sample points in } A}{\text{Number of sample points in } S} = \frac{n(A)}{n(S)} \qquad \ldots (3{\cdot}7)$$

$$= \frac{\text{Number of cases favourable to the event } A}{\text{Exhaustive number of cases}},$$

which is the result obtained by the classical definition of probability.

Hence, *the classical probability may be regarded as a special case of axiomatic probability.*

4. It is interesting to see that there are some formal statements of the properties of events derived from the frequency approach.

From axiomatic definition of probability, we get, $P(A) = \frac{n(A)}{n(S)} \geq 0$, as in Axiom 1.

Also, since $n(S) = N$, we have on using (3·7), $P(S) = \frac{n(S)}{n(S)} = 1$, as in Axiom 2.

Further in the case of two mutually exclusive (or disjoint) events A and B defined by sample points $n(A)$ and $n(B)$, the number of sample points belonging to $A \cup B$ is $n(A) + n(B)$.

$$\therefore \quad P(A \cup B) = \frac{n(A \cup B)}{n(s)} = \frac{n(A) + n(B)}{n(S)}$$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} = P(A) + P(B), \text{ as in Axiom 3.}$$

Hence, the statistical or frequency approach of probability may be regarded as a special case of axiomatic probability.

3·8·6. Algebra of Events. For events A, B, C

(*i*) $A \cup B = \{\omega \in S : \omega \in A \text{ or } \omega \in B\}$ (*ii*) $A \cap B = \{\omega \in S : \omega \in A \text{ and } \omega \in B\}$

(*iii*) $\overline{A}$ (A complement) $= \{\omega \in S : \omega \notin A\}$ (*iv*) $A - B = \{\omega \in S : \omega \in A \text{ but } \omega \notin B\}$

(*v*) Similar generalisations for $\bigcup_{i=1}^{n} A_i$, $\bigcap_{i=1}^{n} A_i$, etc.

(*vi*) $A \subset B \Rightarrow$ for every $\omega \in A$, $\omega \in B$, (*vii*) $B \supset A \Rightarrow A \subset B$.

(*viii*) $A = B$ if and only if A and B have same elements, *i.e.*, if $A \subset B$ and $B \subset A$.

(*ix*) A and B disjoint (mutually exclusive) $\Rightarrow A \cap B = \phi$ (null set)

(*x*) $A \cup B$ can be denoted by $A + B$ if A and B are disjoint.

(*xi*) $A \Delta B$ denotes those ω belonging to exactly one of A and B, *i.e.*,

$$A \Delta B = \overline{A}B \cup A\overline{B} = \overline{A}B + A\overline{B} \text{ (disjoint events)}$$

Remark. Since the events are subsets of S, all the laws of set theory, *viz.*, commutative laws, associative laws, distributive laws, De-Morgan's law, etc., hold for algebra of events.

TABLE – GLOSSARY OF PROBABILITY TERMS

Statement	Meaning in terms of set theory
1. At least one of the events A or B occurs	$\omega \in A \cup B$
2. Both the events A and B occur.	$\omega \in A \cap B$
3. Neither A nor B occurs	$\omega \in \overline{A} \cap \overline{B}$
4. Event A occurs and B does not occur.	$\omega \in A \cap \overline{B}$
5. Exactly one of the events A or B occurs.	$\omega \in A \Delta B$
6. Not more than one of the events A or B occurs.	$\omega \in (A \cap \overline{B}) \cup (\overline{A} \cap B) \cup (\overline{A} \cap \overline{B})$
7. If event A occurs, so does B.	$A \subset B$
8. Events A and B are mutually exclusive.	$A \cap B = \phi$
9. Complementary event of A.	$\overline{A}$
10. Sample space	Universal set S.

Example 3·23. *A, B and C are three arbitrary events. Find expression for the events noted below, in the context of A, B and C :*

(i) Only A occurs, *(ii) Both A and B, but not C, occur,*
(iii) All three events occur, *(iv) At least one occurs,*
(v) At least two occur, *(vi) One and no more occurs,*
(vii) Two and no more occur, *(viii) None occurs.*

Solution.

(i) $A \cap \overline{B} \cap \overline{C}$, *(ii)* $A \cap B \cap \overline{C}$, *(iii)* $A \cap B \cap C$ *(iv)* $A \cup B \cup C$,

(v) $(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) \cup (A \cap B \cap C)$

(vi) $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$

(vii) $(A \cap B \cap \overline{C}) \cup (\overline{A} \cap B \cap C) \cup (A \cap \overline{B} \cap C)$

(viii) $(\overline{A} \cap \overline{B} \cap \overline{C})$ or $\overline{A \cup B \cup C}$

3·9. SOME THEOREMS ON PROBABILITY

In this section, we shall prove a few simple theorems which help us evaluate the probabilities of some complicated events in a rather simple way. In proving these theorems, we shall follow the axiomatic approach, based on the three axioms, discussed in § 3·8·5.

Theorem 3·2. *Probability of the impossible event is zero, i.e., $P(\phi) = 0$.* ... (3·8)

Proof. Impossible event contains no sample point and hence the certain event S and the impossible event ϕ are mutually exclusive.

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

Hence, using Axiom 2 of probability, *i.e.,* Axiom of Additivity, we get

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

Remark. It may be noted $P(A) = 0$, does not imply that A is necessarily an empty set. In practice, probability '0' is assigned to the events which are so rare that they happen only once in a lifetime. For example, if a person who does not know typing is asked to type one page of the manuscript of a book, the probability of the event that he will type it correctly without any mistake is 0.

As another illustration, let us consider the random tossing of a coin. The event that the coin will stand erect on its edge, is assigned the probability 0.

The study of continuous random variable provides another illustration to the fact that $P(A) = 0$, does not imply $A = \phi$, because in case of continuous random variable X, the probability at a point is always zero, *i.e.*, $P(X = c) = 0$ [See Chapter 5].

Theorem 3·3. *Probability of the complementary event $\bar{A}$ of A is given by*

$$P(\bar{A}) = 1 - P(A). \quad \ldots (3\cdot9)$$

Proof. A and $\bar{A}$ are mutually disjoint events, so that

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

Hence, from Axioms 2 and 3 of probability, we have

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

Cor. 1. We have $P(A) = 1 - P(\bar{A}) \leq 1 \quad [\because P(\bar{A}) \geq 0$, by Axiom 1]

Further, since $P(A) \geq 0$ (Axiom 1)

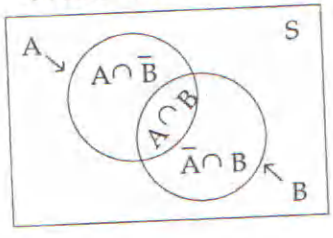
$$\therefore \quad 0 \leq P(A) \leq 1. \quad \ldots (3\cdot9a)$$

Cor. 2. $P(\phi) = 0$, since $\phi = \bar{S}$ and $P(\phi) = P(\bar{S}) = 1 - P(S) = 1 - 1 = 0$.

Theorem 3·4. *For any two events A and B, we have*

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

Proof. From the Venn diagram, we get $B = (A \cap B) \cup (\bar{A} \cap B)$,



where $\bar{A} \cap B$ and $A \cap B$ are disjoint events.

Hence by Axiom (3), we get

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

$$\Rightarrow \quad P(\bar{A} \cap B) = P(B) - P(A \cap B) \quad \ldots (3\cdot10)$$

(*ii*) Similarly, we have $A = (A \cap B) \cup (A \cap \bar{B})$,

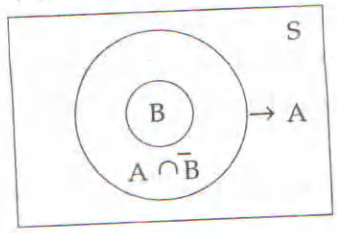
where $(A \cap B)$ and $A \cap \bar{B}$ are disjoint events. Hence, by Axiom 3, we get

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) \quad \Rightarrow \quad P(A \cap \bar{B}) = P(A) - P(A \cap B) \quad \ldots (3\cdot11)$$

Theorem 3·5. *If $B \subset A$, then*

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

(*ii*) $P(B) \leq P(A)$



Proof. (*i*) When $B \subset A$, B and $A \cap \bar{B}$ are mutually exclusive events so that $A = B \cup (A \cap \bar{B})$

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

$$= P(B) + P(A \cap \bar{B}) \quad \text{(By Axiom 3)}$$

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

(*ii*) $P\ (A \cap \bar{B}) \geq 0 \quad \Rightarrow \quad P(A) - P(B) \geq 0 \quad \Rightarrow \quad P(B) \leq P(A)$

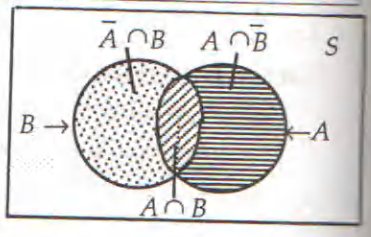
Hence $B \subset A \quad \Rightarrow \quad P(B) \leq P(A)$ ... (3·12)

3·9·1. Addition Theorem of Probability

Theorem 3·6. *If A and B are any two events (subsets of sample space S) and are not disjoint, then*

$$P(A \cup B) = P\ (A) + P\ (B) - P(A \cap B) \qquad \ldots (3·13)$$

Proof. From the Venn diagram, we have



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

where A and $\bar{A} \cap B$ are mutually disjoint.

$$\therefore \quad P(A \cup B) = P[A \cup (\bar{A} \cap B)]$$

$$= P(A) + P(\bar{A} \cap B) \qquad \text{[By Axiom 3)}$$

$$= P(A) + P(B) - P(A \cap B) \qquad \ldots (*)$$

[From Theorem 3·3 (*i*)]

OR From (*) onwards.

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

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

$[\because (\bar{A} \cap B)$ and $(A \cap B)$ are disjoint]

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

Aliter. $P(A \cup B) = \dfrac{n(A \cup B)}{n(S)} = \dfrac{n(A) + n(B) - n(A \cap B)}{n(S)}$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A \cap B)}{n(S)} = P(A) + P(B) - P(A \cap B)$$

Cor. 1. If the events A and B are mutually disjoint, then

$$A \cap B = \phi \quad \Rightarrow \quad P(A \cap B) = P(\phi) = 0$$

$\therefore \quad P(A \cup B) = P(A) + P(B)$, which is Axiom 3 of probability.

Cor. 2. For three non-mutually exclusive events A, B and C, we have

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C). \qquad (3·13a)$$

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

$$= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \qquad \text{[From (3·13)]}$$

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

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

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

3·9·2. Extension of Addition Theorem of Probability to n Events.

Theorem 3·7. *For n events $A_1, A_2, \ldots, A_n$, we have*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum\sum_{1 \le i < j \le n} P(A_i \cap A_j) + \sum\sum\sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k)$$
$$+ \ldots\ldots + (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \quad \ldots (3\cdot14)$$

Proof. For two events A_1 and A_2, we have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \quad (*)$$

Hence (3·14) is true for $n = 2$.

Let us now suppose that (3·14) is true for $n = r$, (say) so that

$$P\left(\bigcup_{i=1}^{r} A_i\right) = \sum_{i=1}^{r} P(A_i) - \sum\sum_{i \le i < j \le r} P(A_i \cap A_j) + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r) \ldots (**)$$

Now

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left\{\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right\}$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left\{\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right)\right\} \quad \ldots \text{[Using (*)]}$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left\{\bigcup_{i=1}^{r} \left(A_i \cap A_{r+1}\right)\right\} \quad \text{(By Distributive Law)}$$

$$= \sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r)$$
$$+ P(A_{r+1}) - P\left\{\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right\} \quad \text{[From (**)]}$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r)$$
$$- \left\{\sum_{i=1}^{r} P(A_i \cap A_{r+1}) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j \cap A_{r+1})\right.$$
$$\left. + \ldots + (-1)^{r-1} P(A_1 \cap A_2 \cap \ldots \cap A_r \cap A_{r+1})\right\} \quad \ldots \text{[From (**)]}$$

$$\Rightarrow P\left(\bigcup_{i=1}^{r+1} A_i\right) = \sum_{i=1}^{r+1} P(A_i) - \left\{\sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \sum_{i=1}^{r} P(A_i \cap A_{r+1})\right\}$$
$$+ \ldots + (-1)^r P\{(A_1 \cap A_2 \cap \ldots \cap A_{r+1}]$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le (r+1)} P(A_i \cap A_j) + \ldots + (-1)^r P(A_1 \cap A_2 \cap \ldots \cap A_{r+1})$$

Hence if (3·14) is true for $n = r$, it is also true for $n = (r + 1)$. But we have proved in (*) that (3·14) is true for $n = 2$. Hence by the principle of mathematical induction, it follows that (3·14) is true for all positive integral values of n.

C : the card drawn is a red card.

Then A, B and C are not mutually exclusive.

$A \cap B$: the card drawn is the king of hearts $\Rightarrow n(A \cap B) = 1$

$B \cap C = B$: the card drawn is a heart $(\because B \subset C) \Rightarrow n(B \cap C) = 13$

$C \cap A$: the card drawn is a red king $\Rightarrow n(C \cap A) = 2$

$A \cap B \cap C = A \cap B$: the card drawn is the king of hearts $\Rightarrow n(A \cap B \cap C) = 1$

$$\therefore \quad P(A) = \frac{n(A)}{n(S)} = \frac{4}{52}; \quad P(B) = \frac{13}{52}; \quad P(C) = \frac{26}{52}$$

$$P(A \cap B) = \frac{1}{52}; \quad P(B \cap C) = \frac{13}{52}; \quad P(C \cap A) = \frac{2}{52}; \quad P(A \cap B \cap C) = \frac{1}{52}$$

The required probability of getting a king or heart or a red card is given by :

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

$$= \frac{4}{52} + \frac{13}{52} + \frac{26}{52} - \frac{1}{52} - \frac{13}{52} - \frac{2}{52} + \frac{1}{52} = \frac{28}{52} = \frac{7}{13}.$$

3·10. CONDITIONAL PROBABILITY

As discussed earlier, the probability $P(A)$ of an event A represents the likelihood that a random experiment will result in an outcome in the set A relative to the sample space S of the random experiment. However, quite often, while evaluating some event probability, we already have some information stemming from the experiment. For example, if we have prior information that the outcome of the random experiment must be in a set B of S, then this information must be used to re-appraise the likelihood that the outcome will also be in B. This re-appraised probability is denoted by $P(A \mid B)$ and is read as the conditional probability of the event A, given that the event B has already happened.

We give below some illustrations to explain this concept.

Illustrations 1. Let us consider a random experiment of drawing a card from a pack of cards. Then the probability of happening of the event A : "The card drawn is a king", is given by : $P(A) = \frac{4}{52} = \frac{1}{13}$.

Now suppose that a card is drawn and we are informed that the drawn card is red. How does this information affect the likelihood of the event A ?

Obviously, if the event B : 'The card drawn is red', has happened, the event 'Black card' is not possible. Hence the probability of the event A must be computed relative to the new sample space 'B' which consists of 26 sample points (red cards only), *i.e.*, $n(B) = 26$. Among these 26 red cards, there are two (red) kings so that $n(A \cap B) = 2$. Hence, the required probability is given by :

$$P(A \mid B) = \frac{n(A \cap B)}{n(B)} = \frac{2}{26} = \frac{1}{13}.$$

2. Consider a random experiment of tossing three fair coins. Then, as explained earlier, the sample space S is :

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

$$= \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}, \text{ (say)},$$

so that : $P(\{\omega_i\}) = \frac{1}{8}; i = 1, 2, \ldots 8.$

Now suppose that the same experiment is performed by another person and thing is known about its outcome. However, we have the information that he tained 'at least two heads'. We are interested to find how this additional rmation affects the probabilities of the elementary outcomes.

This means that if the event A : 'At least two heads are obtained', has happened n the elementary outcomes ω_4, ω_6, ω_7 and ω_8 could not have happened. wever, the remaining four outcomes ω_1, ω_2, ω_3, and ω_5 are still possible and assign the probability $\frac{1}{4}$ to each one of them.

$$\therefore \quad P(\omega_1 \mid A) = P(\omega_2 \mid A) = P(\omega_3 \mid A) = P(\omega_5 \mid A) = \frac{1}{4}.$$

From the above illustrations we observe that some additional information may ange the probability of the happening of some event. We now proceed to develop cedure to calculate the probabilities of events when we know some additional rmation.

Remark. When we know that a particular event B has occurred, instead of S, we centrate our attention on B only and the conditional probability of A given B will be alogously the ratio of the probability of that part of A which is included in B (*i.e.*, $A \cap B$) to probability of B. It, therefore, reflects the change of viewpoint only, namely, instead of S we e to concentrate on B only.

3·11. MULTIPLICATION THEOREM OF PROBABILITY

Theorem 3·9. *For two events A and B,*

$$\left.\begin{aligned} P(A \cap B) &= P(A) \cdot P(B \mid A), \quad P(A) > 0 \\ &= P(B) \cdot P(A \mid B), \quad P(B) > 0 \end{aligned}\right\} \quad \ldots (3\cdot17)$$

where $P(B \mid A)$ represents conditional probability of occurrence of B when the event A has already happened and $P(A \mid B)$ is the conditional probability of happening of A, given that B has already happened.

Proof. In the usual notations, we have

$$P(A) = \frac{n(A)}{n(S)}; \qquad P(B) = \frac{n(B)}{n(S)} \quad \text{and} \quad P(A \cap B) = \frac{n(A \cap B)}{n(S)} \qquad \ldots (*)$$

For the conditional event $A \mid B$, the favourable outcomes must be one of the ample points of B, *i.e.*, for the event $A \mid B$, the sample space is B and out of the (B) sample points, $n(A \cap B)$ pertain to the occurrence of the event A. Hence

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

Rewriting (*), we get

$$P(A \cap B) = \frac{n(B)}{n(S)} \times \frac{n(A \cap B)}{n(B)} = P(B) \cdot P(A \mid B) \qquad \ldots (**)$$

Similarly, we get from (*) :

$$P(A \cap B) = \frac{n(A)}{n(S)} \times \frac{n(A \cap B)}{n(A)} = P(A) \cdot P(B \mid A) \qquad \ldots (***)$$

From (**) and (***), we get the result (3·17).

Thus, we have proved that "*the probability of the simultaneous occurrence of two events*

A and B is equal to the product of the probability of one of these events and the conditional probability of the other, given that the first one has occurred". Any of the events may be called the first event.

Remarks 1. $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$ and $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ ... (3·18)

Thus the conditional probabilities $P(B \mid A)$ and $P(A \mid B)$ are defined if and only if $P(A) \neq 0$ and $P(B) \neq 0$, respectively.

2. (*i*) For $P(B) > 0$, $P(A \mid B) \leq P(A)$

Proof. $n(A \cap B) \leq n(A)$ and $n(B) \leq n(S)$. Dividing, we get

$$\frac{n(A \cap B)}{n(B)} \leq \frac{n(A)}{n(S)} \quad \Rightarrow \quad P(A \mid B) \leq P(A).$$

(*ii*) The conditional probability $P(A \mid B)$ is not defined if $P(B) = 0$.

(*iii*) $P(B \mid B) = 1$.

3·12. INDEPENDENT EVENTS

Two or more events are said to be *independent* if the happening or non-happening of any one of them, does not, in any way, affect the happening of others.

Consider the experiment of throwing two dice, say die 1 and die 2. It is obvious that the occurrence of a certain number of dots on the die 1 has nothing to do with a similar event for the die 2. The two are quite independent of each other, so to say. But suppose, the two dice were connected with a piece of thread before being thrown. The situation changes. This time the two events are not independent in as much as that the uppermost face of one die will have something to do in causing a particular face of the other die to be uppermost ; and the shorter the thread the more is this influence or dependence.

Similarly, if we draw two cards from a pack of cards in succession, then the results of the two draws are independent if the cards are drawn with replacement (*i.e.*, if the first card drawn is placed back in the pack before drawing the second card) and the results of the two draws are not independent if the cards are drawn without replacement.

Definition. *An event A is said to be independent (or statistically independent) of another event B, if the conditional probability of A given B, i.e., $P(A \mid B)$ is equal to the unconditional probability of B, i.e., if* $P(A \mid B) = P(A)$. ... (3·19)

It may be noted that the above definition is meaningful only when $P(A \mid B)$ is defined, *i.e.*, if $P(B) \neq 0$.

Similarly, an event B is said to be independent (or statistically independent) of event A, if

$$P(B \mid A) = P(B); \quad P(A) \neq 0. \qquad \ldots (3·20)$$

Theorem 3·10. *If the events A and B are such that $P(A) \neq 0$, $P(B) \neq 0$ and A is independent of B, then B is independent of A.*

Proof. Since the event A is independent of B, we have

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