UNIT-I

Page No. :
Date: / /

• **Probability:-** According to Ya-Lin Chou, "Probability is the science of decision-making with calculated risks in the face of uncertainty."

• **Basic Terminology:-**

(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.
E.g.→ tossing a coin, throwing a die, selecting a card from a pack of playing cards, etc.

(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.
Eg. In an experiment which consists of the throw of a six-faced die and observing the no. 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 no. of points'; 'Getting a point greater than 4';

Page No. :
Date: / /

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 simple event but the event of getting an even no is a composite event.

(4.) Exhaustive Events or Cases:- The total no. of possible outcomes of a random experiment is known as the exhaustive events or cases. E.g.
(i) In drawing two cards from a pack of cards, the exhaustive no. of cases is ${}^{52}C_2$, since 2 cards can be drawn out of 52 cards in ${}^{52}C_2$ ways.
(ii) In throwing of two dice, the exhaustive no. of cases is $6^2 = 36$, since any of the nos 1 to 6 on the first die can be associated with any of the 6 nos. on the other die. In general, in throwing of n dice, the exhaustive no. of cases is 6^n.

(5.) Favourable Events or Cases:- The no. of cases favourable to an event in a trial is the no. of outcomes which entail the happening of the event.
Eg, In drawing a card from a pack of cards the no. of cases favourable to drawing of an ace is 4, for drawing a spade is 13 and for drawing a red card is 26.

Page No. :
Date: / /

(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.

E.g. (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.

(7.) Equally Likely Events :- Outcomes of trials 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. E.g.

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

(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 occurance of any no. of the remaining events. E.g,

If we draw a card from a pack of well-~~suff~~ 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.

• **Mathematical (or classical or 'A Priori') Probability:-**
If a random experimented or a trial results in 'n' exhaustive, mutually exclusive and equally likely outcomes (or cases), out of which m are favourable to the occurance of an event E, then the probability 'p' of occurance (or happening) of E, usually denoted by P(E), is given by:

$$p = P(E) = \frac{\text{No. of favourable cases}}{\text{Total no. of exhaustive cases}} = \frac{m}{n}$$

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

Page No. :
Date: / /

• Axiomatic Approach to Probability :-

(i) Trial, Random Experiment, Sample point, Outcome & sample space :-

Each performance in a random experiment, is called a trial. That is, all the trials conducted under the same set of condn's form a random experiment. The result of a trial in a random experiment is called an outcome, an elementary event or a sample point. The totality of all possible outcomes (i.e. sample points) of a random experiment constitutes the sample space.

Elementary Event :- The elements or the points of a sample space associated with a random experiment are called elementary events of the experiment.

Sample Space :- The set of all possible outcomes of a given random experiment is called the sample space associated with the experiment. Each possible outcome or element in a sample space is called a sample point or an elementary event.

Page No. :
Date: / /

Discrete and Continuous Sample Space :-

A sample space is called discrete if it contains only finitely or infinitely many points which can be arranged into a simple sequence $\omega_1, \omega_2, \omega_3, \omega_4, \ldots$ while a sample space containing non-denumerable no. of points is called a continuous sample space.

Example of sample space and sample point :-

Consider random tossing of a single coin. The possible outcomes for this experiment are : H & T. Thus the sample space S consists of two points corresponding to each possible outcome or elementary event listed, i.e., $S = \{H, T\}$ and $n(S) = 2$, where $n(S)$ is the total no. of sample points in S.

• **Probability function :-**

$P(A)$ is a probability f^n defined on a σ-field B of events if the following properties or axioms hold.

(i) For each $A \in B$, $P(A)$ is defined, is real and $P(A) \geq 0$ (Axiom of non-negativity)

(ii) $P(S) = 1$ (Axiom of certainty)

(iii) If $\{A_n\}$ is any finite or infinite sequence of disjoint ~~ev~~ events in B, then

Page No. :
Date: / /

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P(A_i)$$ (axiom of additivity)

- **Glossary of Probability Terms:-**

Statement	Meaning in terms of set theory
(i) At least one of the events A or B occurs.	$\omega \in A \cup B$
(ii) Both the events A and B occur.	$\omega \in A \cap B$
(iii) Neither A nor B occurs	$\omega \in \overline{A} \cap \overline{B}$
(iv) Event A occurs and B doesn't occur.	$\omega \in A \cap \overline{B}$

$A \Delta B \rightarrow$ Symmetric Difference means objects that belong to A or B but not to their intersection. Eg $A = \{$[illegible]$4\}$, $B = \{1, 2, 3\}$
$A \Delta B = \{1, 3, 4\}$

(v) Exactly one of the events A or B occurs.	$\omega \in A \Delta B$ since, $A \Delta B = \bar{A}B \cup A\bar{B}$
(vi) Not more than one of the events A or B occurs.	$\omega \in (A \cap \bar{B}) \cup (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B})$
(vii) If event A occurs, so does B.	$A \subset B$
(viii) Events A and B are mutually exclusive	$A \cap B = \phi$
(ix) Complementary event of A	$\bar{A}$
(x) Sample space	Universal set S.

- **Some Theorems On Probability :-**

Thm → Probability of the impossible event is zero, i.e., $P(\phi) = 0$

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

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

Hence, using Axiom of additivity, we get
$$P(S) + P(\phi) = P(S) \Rightarrow P(\phi) = 0.$$

Note :- It may be noted $P(A) = 0$, doesn't 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

Page No. :
Date: / /

E.g. → ① If a person who doesn't know typing is asked to type on page of the manuscript of a book, the probability of the event that he will type it correctly without any mistake is '0'.

② 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'.

Thm:- Probability of the complementary event $\bar{A}$ of A is given by $P(\bar{A}) = 1 - P(A)$

Pf:-

A and $\bar{A}$ are mutually disjoint events, so that $A \cup \bar{A} = S \Rightarrow P(A \cup \bar{A}) = P(S)$

Hence, from axioms of additivity ~~or~~ and certainty of ~~events~~ probability, we have

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

Corollary 1:- We have

$$P(A) = 1 - P(\bar{A}) \leq 1$$ [$\because P(\bar{A}) \geq 0$ by axiom of non-negativity]

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

$$\therefore \quad 0 \leq P(A) \leq 1$$

Corollary 2:- $P(\phi) = 0$, since, $\phi = \bar{S}$ and

$$P(\phi) = P(\bar{S}) = 1 - P(S) = 1 - 1 = 0$$

• **Thm:- Addition Theorem of Probability:-**

Statement:- If A and B are any two events (subsets of sample space S) and are not disjoint then

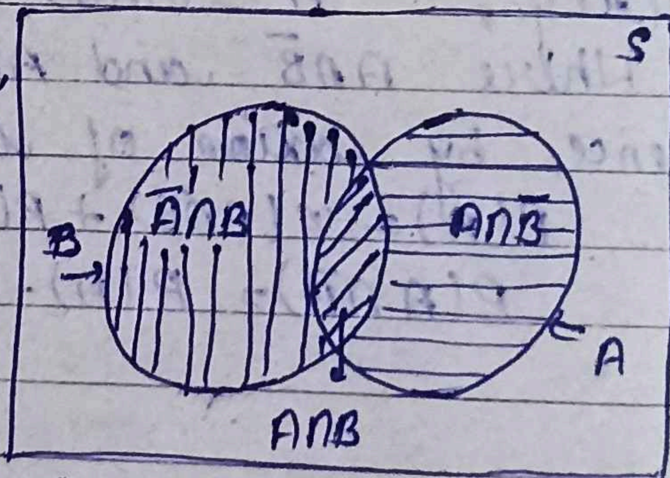
$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

Pf:-

From Venn diagram, we have

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

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



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

$$= P(A) + P(\bar{A}\cap B) \quad \text{[By axiom of tivity]}$$

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

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

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

Corollary: 1 →.

If the events A and B are mutually disjoint, then

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

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

which is axiom of +vity of probability.

Corollary 2:-

For three non-mutually exclusive events A, B & C, we have

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

Pf:-

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

Conditional Probability :-

The probability for the event A to occur when it is known that the event B has already occurred, is called the conditional probability A given B and is denoted by $P(A/B)$ or $P\left(\frac{A}{B}\right)$ and defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad , \text{ provided } P(B) \neq 0$$

Theorem on Compound Probability :-

If A and B are two events in S,

then $$P(A \cap B) = \begin{cases} P(A)\, P(B/A) & , P(A) \neq 0 \\ P(B)\, P(A/B) & , P(B) \neq 0 \end{cases}$$

• **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.

Def:- 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., $P(A/B)$ is equal to the unconditional probability of A, i.e., if $P(A/B) = P(A)$.

Similarly,

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

$$P(B/A) = P(B) \; ; \; P(A) \neq 0.$$

Thm:- 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.

- **Multiplication Theorem of Probability for Independent Events :-**

Thm:- If A and B are two events with +ve probabilities $\{P(A) \neq 0, P(B) \neq 0\}$, then A and B are independent iff $P(A \cap B) = P(A) . P(B)$ — (#)

Pf:- We have :

$$\left.\begin{array}{l} P(A \cap B) = P(A) . P(B|A) = P(B) P(A|B) ; \\ \quad P(A) \neq 0 , P(B) \neq 0 \end{array}\right\} — \text{I}$$

If A and B are independent, i.e., A is independent of B and B is independent of A, then we have

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B) \quad — \text{II}$$

from I & II, we get

$P(A \cap B) = P(A) . P(B)$, as required.

Page No.:
Date: / /

Conversely, if (#) holds, then we get

$$\left.\begin{array}{l} \frac{P(A\cap B)}{P(B)} = P(A) \Rightarrow P(A|B) = P(A) \\ \text{and } \frac{P(A\cap B)}{P(A)} = P(B) \Rightarrow P(B|A) = P(B) \end{array}\right\} \text{— (III)}$$

(III) implies that A and B are independent events.

Hence, for independent events A and B, the probability that both of these occur simultaneously is the product of their respective probability. This rule is known as the Multiplication Rule of Probability.

Bayes' Theorem:-

Bayes' theorem which was given by Thomas Bayes, a British Mathematician, in 1763, provides a means for making these probability calculations. The step in this probability revision process are showed in the following diagram:-

Statement :-

If $E_1, E_2, E_3, \ldots, E_n$ are mutually disjoint events with $P(E_i) \neq 0$, $(i = 1, 2, \ldots, n)$, then for any arbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ such that $P(A) > 0$, we have

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

Proof →

Since, $A \subset \bigcup_{i=1}^{n} E_i$, we have,

$$A = A \cap \left(\bigcup_{i=1}^{n} E_i \right) = \bigcup_{i=1}^{n} (A \cap E_i)$$

[by distributive Law]

Since $(A \cap E_i) \subset E_i$, $(i = 1, 2, \ldots, n)$ are mutually disjoint events, we have by addition theorem of probability:

$$P(A) = P\left\{ \bigcup_{i=1}^{n} (A \cap E_i) \right\} = \sum_{i=1}^{n} P(A \cap E_i)$$

$$= \sum_{i=1}^{n} P(E_i)\, P(A|E_i) \quad (*)$$

by multiplication theorem of probability.
Also we have $P(A \cap E_i) = P(A)\, P(E_i|A)$

$$\Rightarrow \quad P(E_i|A) = \frac{P(A \cap E_i)}{P(A)}$$

[By conditional Probability]

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

(from (*))

Hence Proved

Distribution f^n d.f.

ORI ⑫ II ⑫ III — Random Variables R.V

Page No.:
Date: / /

• Random Variable:-

By a random variable we mean a real no. X connected with the out-come of a random experiment E. For example, if E consists of two tosses the random variable which is the no. of heads (0, 1, or 2).

Outcome :	HH	HT	TH	TT
Value of X :	2	1	1	0

Thus to each outcome ω, there corresponds a real no. $X(\omega)$. Since the points of the sample space S correspond to outcomes, this means that a real no, which we denote by $X(\omega)$, is defined for each $\omega \in S$. From this standpoint, we define random variable to be a real f^n on S as follows:

"Let S be the sample space associated with a given random experiment. A real valued f^n defined on S and taking values in $R(-\infty, \infty)$ is called a one-dimensional random variable. If the f^n values are ordered pairs of real no.; the f^n is said to be a two-dimensional random variable. More generally, an n-dimensional random variable is simply a f^n whose domain is S and whose range is a collection of n-tuples of real no.'s"

Distribution function:-

Let X be a random variable. The f^n F defined for all real x by

$$F(x) = P(X \leq x) = P\{\omega : X(\omega) \leq x\}, -\infty < x < \infty,$$

is called the distribution f^n of the r.v. X.

Let for the previous eg, $P(X=0) = 1/4$, $P(X=1) = 2/4 = 1/2$, $P(X=2) = 1/4$

Remark:- A distribution f^n is also called the cumulative d.f.. Clearly, the domain of the d.f. is $(-\infty, \infty)$ and its range is $[0, 1]$.

Properties of D.F. →

(i) If F is the d.f. of the r.v. X and if $a < b$, then $P(a < X \leq b) = F(b) - F(a)$

(ii) If F is d.f. of one-dimensional r.v. X, then (i) $0 \leq F(x) \leq 1$ (ii) $F(x) \leq F(y)$ if $x < y$.
In other words, all the distribution f^n's are monotonically non-decreasing and lies b/w 0 and 1.

(iii) If F is d.f. of one-dimensional r.v. X, then

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0 \text{ and}$$

$$F(\infty) = \lim_{x \to \infty} F(x) = 1$$

Page No. :
Date: / /

- ## Discrete Random Variable :-

A variable which can assume only a countable no. of real values and for which the values ~~which~~ that the variable takes depends on chance, is called a discrete random variable.

In other words, a real valued f^n defined on a discrete sample space is called a d.r.v.

E.g., (i) marks obtained in a test (ii) no. of accidents per month (iii) no. of telephone calls per unit time, no. of successes in n trials and so on.

Probability Mass f^n / ~~Discrete density f^n~~ :-

If X is a one-dimensional d.r.v. taking at most a countably infinite no. of values $x_1, x_2, \ldots$ ... then its probabilistic behaviour at each real point is described by a f^n called the P.M.f^n (or d.d.f.) which is defined below:-

Def:- If X is a d.r.v. with distinct values $x_1, x_2, \ldots x_n, \ldots$, then the f^n $p(x)$ defined as:

$$p_X(x) = \begin{cases} P(X = x_i) = p_i & \text{if } x = x_i \\ 0 & \text{if } x \neq x_i \,;\, i = 1, 2, \ldots \end{cases}$$

is called the p.m.f^n of r.v. X.

Page No. :
Date: / /

The set of ordered pairs $\{x_i, p(x_i) ; i = 1, 2, \ldots, n, \ldots\}$ or $\{(x_1, p_1), (x_2, p_2), \ldots, (x_n, p_n), \ldots\}$ specifies the probability distribution of the r.v. X.

E.g. Toss of a coin, $S = \{H, T\}$. Let X be a r.v. defined by:

$X(H) = 1$, i.e., $X = 1$ if 'Head' occurs.
$X(T) = 0$, i.e., $X = 0$ if 'Tail' occurs.

If the coin is 'fair'; the probability f^n is

$$P(\{H\}) = P(\{T\}) = \frac{1}{2},$$

and we can speak of the probability distribution of the r.v. X as:

$$P(X = 1) = P(\{H\}) = \frac{1}{2} \; ; \text{ and } P(X = 0) = P(\{T\}) = \frac{1}{2}.$$

• **Continuous Random Variable :-** A random variable X is said to be continuous if it can take all possible values (integral as well as fractional) b/w certain limits.

In other words, a r.v. is said to be continuous when its different values can't be put in 1-1 correspondence with a set of +ve integers.

A C.R.V. is a r.v. that (at least conceptually) can be measured to any desired degree of accuracy.

Examples of C.R.V. are age, height, weight, etc.

Probability Density Function →

Let the probability that the continuous random variable X lies within infinitsmall interval $\left(x-\frac{dx}{2}, x+\frac{dx}{2}\right)$ is

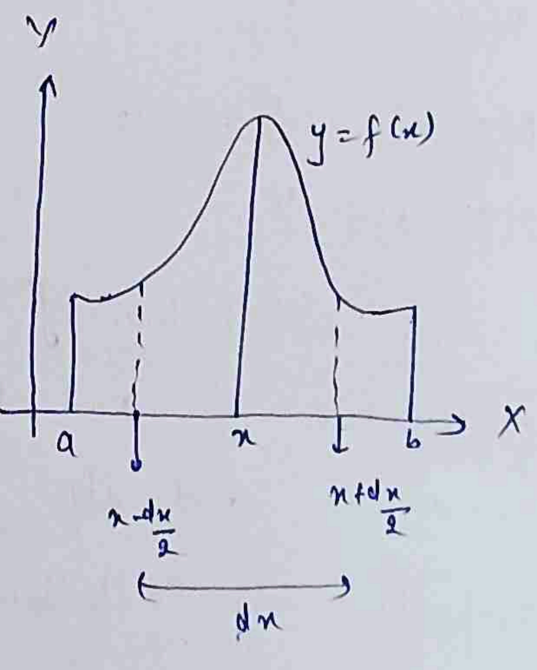
$$P\left(\left(x-\frac{dx}{2}\right) \leq X \leq \left(x+\frac{dx}{2}\right)\right) = f(x)\,dx$$

where $f(x)$ is a continuous f^n of x, is called the probability density f^n and the curve $y = f(x)$ is called the probability curve.

If the interval is $[a, b]$

then

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



Mathematical Expectation :-

(i) For Discrete Random Variable :-

If X is a discrete random variable which takes values x_i with probability $P(X = x_i) = p_i$, $i = 1, 2, ---$, then the mathematical expectation of X, denoted by $E(X)$ and is defined by

$$E(X) = \sum_{i=1} x_i p_i$$

(ii) For Continuous Random Variable :-

If X is a continuous random variable with probability density function $f(x)$, then the expected value of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Properties of Expectation :-

(i) **Additional Property** :- If X and Y are two random variables, then $E(X+Y) = E(X) + E(Y)$

(ii) **Linear Property** :- If X and Y are two random variables and 'a' and 'b' are constants, then

$$E(aX + bY) = aE(X) + bE(Y)$$

(iii) **Multiplication Property** :- If X and Y are two independent random variables then

$$E(X \cdot Y) = E(X)\, E(Y).$$