UNIT I

INTRODUCTION TO PROBABILITY

Probability is the study of random experiments. In random experiment there is always an uncertainty as to whether a particular event will occur or not. As a measure of chance of probability of an event, a number between 0 and 1 is assigned.

1. If we are sure that the event will occur, then we say its probability is 100% or 1.

2. If we are sure that the event will not occur, then we say its probability is 0% or 0.

3. If we are not sure that whether the probability will occur or not, then its probability is between 0 and 1.

Example:

1. The probability of occurrence of 28 February in a year is 1, as it is certain to occur every year.

2. The probability of occurrence of 30 February in a year is 0, as it never occurs.

3. The probability of occurrence of 29 February in a year is neither 0 nor 1. Its between 0 and 1(It is ¼ as it occurs every leap year).

Three approaches to understand probability

1. Classical or Priori approach

2. Frequency or a Porteriori approach

3. Axiomatic approach.

CLASSICAL OR A PRIORI APPROACH

If an event A can occur in S different ways out of a total of n equally likely sample ways then the probability of an event A is given by,

$$P=P(A)=s/n$$

FREQUENCY APPROACH

If in n repetitions of an experiment where n is large, an event A occurs s times, then the probability of the event is given by $p=P(A) = s/n$.

The ratio s/n, called the relative frequency, becomes stable as n increases.

AXIOMATIC APPROACH

Let s be the sample space consisting of all possible outcomes of an experiment. The events A,B,C.....are subsets of S. Then the function P(.) has to satisfy the following axioms.

Axiom 1: For every event A,$0\leq P(A)\leq 1$

Axiom 2: For sure or certain event $P(S)=1$

Axiom 3: If A and B are mutually exclusive events ie., if $AB=0$ then $P(A+B)=P(A)+P(B)$

Theorem 1

If ϕ is an empty set then $P(\phi) = 0$

$$P(A+\phi)=P(A)$$

$P(A)+P(\phi)=P(A)$,subtracting $P(A)$ on both the sides then

$$P(\phi)=0$$

Theorem 2

$$P(\acute{A})=1-P(A)$$

Theorem 3

If $A<B$ then $P(A)\leq P(B)$

Theorem 4

If A and B are two events then

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

Theorem 5

If A and B are any two events, then

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

CONDITIONAL PROBABILITY

Conditional probability is the probability of some event A, given the occurrence of some other event B. It is written as $P(A/B)$ read as the probability of A given B.

JOINT PROBABILITY

Joint probability is the probability of two events in conjunction. That it is the probability of both the events together. Joint probability of A and B is given as $P(A,B)$.

CONDITIONAL PROBABILITY AND STATISTICAL INDEPENDENCE

Let us consider that drawing of two cards as a single event, then let drawing of diamond be a event D, and drawing of heart be a event H.

The $P(DH)$ is the probability of drawing first a diamond and then a heart. $P(HD)$ is the probability of drawing first a heart and then a diamond. $P(DH)$ and $P(HD)$ are known as joint probabilities, then desired probability is $P = P(DH)+P(HD)$

The probability of drawing diamond as the first card, $P(D)=13/52 =1/4$.Since 13 out of the 52 cards are diamonds.

The probability of drawing a heart as the second card is 13/51, as only 51 cards are left. Thus the second drawing is dependent upon or conditioned by the first drawing.

The probability corresponding to the second drawing is designated by the symbol $P(H/D)$ is known as conditional probability .

Note: If first is replaced then the probability of second drawing is independent of first drawing, then $P(H)=13/52=1/4$.

The probability of drawing first as diamond and then heart is

$$P(DH)=P(D)P(H/D)=(13/52)(13/51)$$

First heart and then diamond,

$$P(HD)=P(H)P(D/H)=(13/52)(13/51)$$

$$P=P(DH)+P(HD)$$

$$=(13/52)(13/51)+ (13/52)(13/51)=13/102.$$

JOINT AND CONDITIONAL PROBABILITY

We can now generalize the concepts of joint probability And conditional probability.Let n be the number of times an experiment is performed. Let n be a large number then the outcome A appears n_A times.The outcome B appears n_B times and combination AB appears n_{AB} times.

Since n is a large number, the joint probability Of first A and then B occuring is

$$P(AB)= n_{AB}/n.$$

$$P(AB)=n_{AB}/n_A$$

$$P(A)=n_A/n,P(B)=nb/n$$

$$P(B/A)=(n_{AB}/n)/(n_A/n)=P(AB)/P(A)$$

$$P(AB)=P(A)P(B/A)$$

$$P(BA)=n_{BA}/n \text{ then } P(A/B)=n_{BA}/n_B$$

$=(n_{BA}/n)/n_B/n=P(BA)/P(B)$. This is invalid when $P(A)$ and $P(B)$ are zero.

Therefore $P(B/A)=P(AB)/P(B),P(A)\neq 0$

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

Now let us consider a situation where the probability of the event B occurring is independent of the event A. Such a situation would be true in two card problem. If the first card were immediately replaced after having been drawn.$P(B/A)=P(B)$.

$$P(AB)=P(A)P(B), \text{similarly } P(A/B)=P(A), \text{ then}$$

$$P(BA)=P(B)P(A), \text{ then } P(AB)=P(BA)=P(A)P(B)$$

RANDOM VARIABLE

Random variable is a real valued function defined over the sample space of a random experiment. It is also known as stochastic variable or stochastic function or random function. The random variables are denoted by upper case letters such as X,Y,etc. And the values assumed by them are denoted by lower case letters with subscripts such as x_1,x_2,y_1,y_2 etc.

There are two types of random variables

1. Discrete random variable
2. Continuous random variable

DISCRETE RANDOM VARIABLE

A random variable that takes on a finite number of values is known as discrete random variable.

Discrete Probability distribution:

Let X be a discrete random variable and let x_1,x_2,x_3...be the values that can assume in the increasing order of magnitude.

Let $P(X=x_j)=f(x_j), j=1,2,3....$ be the probability of x_j. Let there be a function f(x) such that

1. $F(x)\geq 0$
2. $\sum_x f(x) = 1$

Then f(x) is known as probability distribution function.

Cumulative Distribution function for a discrete random variable:

The cumulative distribution function or distribution function of a discrete random variable is defined as

$$F(X)= P(X\leq x)=\sum_{u\leq x} f(u), \quad -\infty<x<\infty$$

If X can take on the values $x_1,x_2,x_3.....x_n$ then the distribution function is given by

$F(x)= \quad 0,$	$-\infty < x < x_1$
$F(x_1),$	$x_1 < x < x_2$
$F(x_1)+f(x_2),$	$x_2 < x < x_3$
$F(x_1)+.....+f(x_n)$	$x_n < x < \infty$

CONTINUOUS RANDOM VARIABLE

A random variable that takes on an infinite number of values is known as continuous random variable. As there are infinite number of possible values of X, the probability that it takes on any particular value is 0. Hence probability function in this case cannot be defined as in the discrete case.

Continuous Probability Distribution:

Let there be a function f(x) such that

$$F(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

Here the function f(x) is known as probability function or probability distribution for a continuous random variable, but it is more popularly known as probability density function.

The probability of X lying between a and b is defined by, P(a< X<b) $=\int_a^b f(x)\,dx$.

BIONOMIAL DISTRIBUTION

The probability of x successes in n trails is given by a probability function known as bionomial distribution;

$$F(x)=P(X=x)= \begin{matrix} n \\ x \end{matrix} \; p^x q^{n-x}$$

Where mean = np ; variance = npq.

POISSON DISTRIBUTION

Let X be a discrete random variable that can assume values 0,1,2.....Then the probability function of X is given by poisson distribution.

$$F(x)=P(X=x)= \lambda^x e^{-\lambda}/x! \text{ where } x=0,1,2.....$$

Where λ is a positive constant, λ = variance=mean.

RAYLEIGH DISTRIBUTION

F(x) = $x/a^2.e^{-\frac{r^2}{2a^2}}$ for $0\leq r \leq \infty$

0 for a>0.

Attains maximum at r=a

$$\mu = a(\frac{\sqrt{\pi}}{2}), \sigma^2 = 2a^2$$

NORMAL OR GAUSSIAN DISTRIBUTION

This is the most important continuous probability distribution as most of the natural phenomenon are characterized by random variables with normal distribution.The density function of normal or Gaussian distribution is

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}.\frac{e^{-(x-\mu)^2}}{2\sigma^2}, -\infty < x < \infty$$

Where μ and σ $a$re mean and standard deviation. The corresponding distribution function is given by

$F(x)=P(X\leq x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-(x-\mu)^2}}{2\sigma^2}dx \quad = ½+\frac{1}{\sigma\sqrt{2\pi}}\int_{0}^{x}\frac{e^{-(x-\mu)^2}}{2\sigma^2}dx$

$F(z)=1/2\{1+erf(z/(\sqrt{2})\} = 1 - 1/2erfc(z/\sqrt{2})$

STATISTICAL AVERAGE

It is to be noted that since in communication theory we have to deal with statistical quantities and not with deterministic quantities, wherever average is referred it always means statistical average and not arithmetic average. Let us apply the definition to a problem, for an example Let M be a transmitted message and the information associated with jth message be Ixj where j=1,2,....Then according to arithmetic mean the average of transmitted message is

$I'x_j=\sum_{j=1}^{M} Ixj\,/M$.....But this definition is not correct.The reason is simple, the transmitted messages are not transmitted once.

CHARASTERISTIC FUNCTION

Expectation:

Let X be a random variable such that [X]=[x1,x2....xm]

$$\mu x = E(x) = \sum_{j=1}^{m} xj/(m)$$

$$\mu x = \int_{-\infty}^{\infty} xf(x)dx$$

Variance:

$$Var(x)=\sigma x^2 = E[(X-\mu)^2]$$

Standard Deviation:

$$\sigma x = \sqrt{Var(x)}$$

CENTRAL LIMIT THEOREM

Central limit theorem states that the probability density of a sum of N independent random variables tends to approach a normal density as the number N increases. The mean and variance of this normal density are the sums of mean and variance of N independent random variables.

Mean $\sigma i = \sigma$ and variance $\sigma i^2 = \sigma^2/n$

If X1,X2,...Xn is a sequence of n Independent and identically distributed random variables, each having mean μ *and variance* σ^2 and if X′=(X1+X2+...+Xn)/n, then the variance Z=(X′- $\mu)/\sigma\sqrt{n}$ has a distribution that approaches the standard normal distribution as n→ ∞.

M.G.F of Z is given by

$M_Z(t)=E[e^{tZ}]$

$=E(e^{t(X-\mu)}/\sigma\sqrt{n})$

E(X1....Xn)=E(X1)E(X2).....E(Xn)

$M_z(t)=e^{-t\mu\sqrt{n}/\sigma}[M_X(t/(\sigma\sqrt{n}))]^n$

Taking log on both the sides

$\log M_Z(t)=-t\frac{\mu\sqrt{n}}{\sigma}+[(\frac{t}{\sigma\sqrt{n}}\mu 1+\frac{\mu 2}{!}(\frac{t}{\sigma\sqrt{n}})^2+......)-1/2(\frac{\mu 1t}{\sigma\sqrt{n}}+\cdots)^2$

$\log M_Z(t)=-t\frac{\mu\sqrt{n}}{\sigma}+\frac{\sqrt{n}\mu t}{\sigma}+t^2/2\sigma^2[\mu 2-(\mu 1)^2]$

$=t^2/2\sigma^2.\sigma^2$

$M_z(t)=e^{t2/2}$ as n→ ∞.