Unit-16:Probability Distributions and Their Applications

16.0 : Objectives

After reading this Unit, you will be able to:

- understand the basic concepts of probability;
- find the probability of occurrence of an event;
- understand the probability distribution of a random variable; and
- learn about the joint probability and marginal probability distribute

16.1 : Introduction

A researcher makes observations resulting in generation of data and information. Data generally demands an inference to be drawn and statistical inferences are never certain as in deterministic experiments. Statisticians tend to use prefixes or suffixes as "likely" or "almost likely". Understanding of probability ad its concepts is, thus, very important to interpret inferential statistics. In a n adorn experiment set all possible outcomes may be known but it is not possible to predetermine the outcome. For example, the probability of a coin toss resulting in either "heads" or "tails" is 1. The probability of a coin toss resulting in "heads" is 0.5, because the toss is equally as likely to result in "tail " and there are only two possible outcomes. Thus, probability is a branch of statistics that deals with calculating the likelihood of a given event's occurrence, whi*ch* is expressed as a number between 1 and O. An event with a probability' of 1can be considered a certain event and an event with a probability of 0 can be considered as an impossible event. In other words, with the help of probability theory you would be equipped with the methods to assign a quantitative measure to the uncertainty associated with the occurrence of each possible outcome of a random experiment.

16.2 : Probability - Definition

Probability is defined in many ways. We would restrict ourselves to two definitions namely classical definition and the axiomatic definition. Before defining probability, definitions of several other terms are needed.

Random Experiment: An experiment or trial whose outcome is not perfectly predictable, but for which the set of all possible outcomes in repeated trials is known.
Random Variable: A random variable is an assignment of numbers to possible outcome of a random experiment. For example, consider tossing two coins. The number of heads showing when the coins land is a random variable. It assigns the number 0 to the outcome {T, T}, the number 1 to the outcome {T, H}, the number 2 to the outcome {H, H}
Outcome Space: The outcome space is the set of all possible outcomes of a given random experiment. The

outcome space is often denoted by the capital letter **S.**

Event: An event is a subset of outcome space. An event is determined by a random variable. Outcomes are the set of all possible results of a Random Experiment.

Exhaustive Events: A collection of events {A,, A,, A_3, . . .} is exhaustive if at least one of them must occur; that is, if S = A, U A_2 U A_3 U . . . where S is the outcome space.

Disjoint or Mutually Exclusive Events: Two events are disjoint or mutually exclusive if the occurrence of one is incompatible with the occurrence of the other; that is, if they cannot both happen at the same time (if they have no outcome in common). Equally, two events are disjoint if their intersection is the empty set. ***Independent Events:*** *Two* events A and B are statistically independent if the chance that they both happen simultaneously is the product of the chances that each occurs individually. In other words, the knowledge that one event has occurred does not give any information about whether the other event has occurred too. ***Equally Likely Events:*** *Two* events are said to be equally likely if the occurrence of none of them is expected in preference to others.

Now, let us define probability. The definition given here is also called mathematical definition of probability given by Bernoulli.

If n is the number of equally likely, mutually exclusive and exhaustive outcomes of a random experiment out of which m outcomes are favourable to the occurrence of an event A, then the probability that A occurs, denoted by P(A), is given by:

$$P(A) = \frac{\text{Number of outcomes favourable to } A}{\text{Number of exhaustive outcomes}} = \frac{m}{n}$$

Before defining the probability by the axiomatic approach, it is important that we review some concepts of set theory, which will enable better understanding of the concept of probability.

16.2.1 : Set Theory Definitions

Set: A set is a collection of things, without regard to their order. We will denote a set with a capital Letter; A, B etc.

An Element of a Set: The principle primitive concept of set theory is that of *belonging. A* number x is an element of A and is denoted by $\mathbf{x \in A}$

Subset: If *A* and B are sets, and every element of A is an element of B, we call A, a subset of B, $A \subset B$.

Empty Set: The empty set, f, is a set that contains no elements.

Complement of a Set: *If A* is a subset of B, then the complement of A relative to B, A' is the subset of B that contains no element of A.

Union of Sets: *Any* two sets may be joined together to form a new set, $A \cup B$ which is a set whose elements are either in A or in B**.**

Intersection of Sets: The intersection of two sets A and B, is a set whose elements are in both A and B, $A \cap B$.

16.2.2 : Axioms and Theorems of Probability

In axiometric approach, probability is introduced as a function of outcomes of an experiment, under certain restrictions. These restrictions are called Axioms of Probability.

Axiom 1: The probability *of* an event is a non-negative real number; that is, $0 \leq P(A) \leq 1$ for any subset *A* of S.

Axiom 2: $P(S) = 1$, i.e. the probability of outcome space is 1

Axiom 3: If A_1, A_2, A_3......... is a finite or infinite sequence of mutually exclusive events of S, then

$$P(A_1 \cup A_2 \cup A_3 \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

The first axiom implies that the probability of an event is a non-negative number less than or equal to unity. The second axiom implies that the probability of an event that is certain to occur must be equal to unity and the last axiom gives a basic rule of addition of probabilities.

How do we assign probabilities to events? There are a few theorems on probability, which will help you in mathematical treatment of the subject of probability.

Theorem 1: $P(\phi)=0$, where ϕ is a null event.

Theorem 2: $P(\overline{A}) = 1 - P(A)$ where $\overline{A}$ is compliment of A.

Theorem 3: For any two events A and B in a sample space S

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

Theorem 4: *Addition of Probabilities*

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

Theorem 5: *Conditional Probability*

For any two events in a sample space S, the probability of their simultaneous occurrence is given by

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B).$$

Where P(B|A) is the conditional probability of B given that A has already occurred.

Theorem 6: Multiplicative *Theorem for Independent Events*

If A and B are independent, the probability of their occurrence simultaneously is given by

$$P(A \cap B) = P(A)P(B).$$

Theorem 7: *Bayes' Theorem*

Before stating the theorem it is necessary to know about *prior* and *posterior probabilities.*

Prior probabilities: These are the probabilities, assigned to events on the basis of conditions, past experience or judgment.

Posterior probabilities: These are the revised probabilities in the light of some additional information.

Bayes' theorem states that if an event B can occur in combination with any of the n mutually exclusive and exhaustive events A_i's and if B is found to have occurred, then the probability that it was preceded by a particular event A_k is given by

$$P(A_K|B) = \frac{P(A_K)P(B|A_K)}{\sum_{I=1}^{n} P(A_I)P(B|A_I)}$$

Let us consider an example.

Example 1 The contents of urns I, II & III are as follows :

1 white , 2 black and 3 red balls

2 white , 1 black and 1 red balls

4 white , 5 black and 3 red balls

one urn is chosen at random and two balls drawn. They happen to be white and red. What is the probability that they come from urn II?

Answer

Let E1,E2 and E3 denote the events that the urn I, II and III is chosen, respectively, and let A be the event that the two balls taken from the selected urn are white and red. Then

$P(E1)=P(E2)=P(E3)=1/3$

$$P(A|E1) = \frac{1\times 3}{{}^{6}C_2} = \frac{1}{5} \quad P(A|E2) = \frac{(2\times 1)}{{}^{4}C_2} = \frac{1}{3} \quad \& \quad P(A|E3)= \frac{(4\times 3)}{{}^{12}C_2} = \frac{2}{11}$$

Hence,

$$P(E_2|A) = \frac{P(E2).P(A|E2)}{\sum_1^3 P(Ei).P(A|Ei)} = \frac{\frac{1}{3}.\frac{1}{3}}{\frac{1}{3}.\frac{1}{5}+\frac{1}{3}.\frac{1}{3}+\frac{1}{3}.\frac{2}{11}} = \frac{55}{1}$$

16.4 : Joint Probability Distribution

When two random variables are studied together then we get a joint probability distribution. Let the random variable X take value Xi and random variable Y take value Yj If p_{ij} be the joint probability that

X takes the value x; an Y takes the value x_i then

	Y_1	Y_2	...	...	Y_n	Marginal Probability of X
X_1	p_{11}	p_{12}	...	...	p_{1n}	P_1
X_2	P_{21}	p_{22}	...	...	p_{2n}	P_2
.			...	...		
.			...	...		
X_m	p_{m1}	p_{m2}	...	...	p_{mn}	P_m
Marginal Probability of Y	P_1'	P_2'	...	...	P_n'	1

Here the probabilities given in each row and each column are added and these are called marginal probabilities for various values of random variable X. he marginal probability distribution of the random variable is represented as he set of all possible values of the random variable along with the marginal probabilities. We can define.

$$p_{ij} = P(X = X_i, Y = Y_j) = P(X = X_i).P(Y = Y_j) \forall i \,\&\, j$$

16.5 : Conditional Probability Distribution

The conditional distribution of X given Y is given by the following formula P(X=Xi | Y=Y_j) = Joint Probability of Xi and Y_j / Marginal Probability of

$$Y_i = \frac{p_{ij}}{P_j^{'}}$$

The conditional distribution of X given Y=Y_1 can be written in tabular form as follows

X	X_1	X_2	...	X_m	Total Probability
Probability	$\frac{P_{11}}{P_1}$	$\frac{P_{21}}{P_1}$	...	$\frac{P_{m1}}{P_1}$	1

Example 3

Given the following bivariate probability distribution, obtain the marginal distribution of X and Y and the conditional distribution of X given Y=2

Y\X	-1	0	1
0	1/15	2/15	1/15
1	3/15	2/15	1/15
2	2/15	1/15	2/15

Answer

Y\X	-1	0	1	p(x,y)
0	1/15	2/15	1/15	4/15
1	3/15	2/15	1/15	6/15
2	2/15	1/15	2/15	5/15
p(x,y)	6/15	5/15	4/15	1

Marginal distribution of X can be seen from the above table as

P(X=-1)= 6/15

P(X=0)=5/15

P(X=1)=4/15

Marginal distribution of Y can also be obtained from the above table as

P(Y=0)= 4/15

P(Y=1)=6/15

P(Y=2)=5/1 5

For conditional distribution of X given Y=2

We have $P(X=x \cap Y=2)=P(Y=2).P(X=x|Y=2)$

So

$P(X=x|Y=2)= P(X=x \cap Y=2)/P(Y=2)$.

Therefore,

$P(X=-1|Y=2)= P(X= -1 \cap Y=2)/P(Y=2) =(2/15)/(1/3)= 2/5$

16.6 : Some Special Distributions

In this section we will give two discrete probability distributions and continuous distribution. These are special distributions and have man applications in statistics. There are many more interesting distributions.

16.6.1 : Binomial Distribution

If an experiment consists of a finite number of independent repeated trial; where each trial has only two possible mutually exclusive outcomes called "success" (probability of success is "p") or "failure" (probability of failure is 1-p), then these trials are called Bernoulli trials. The number of "successes" in n independent random trials is a random variable with Binomial Distribution (with parameters n and p). Under these assumptions, the probability of x

successes is $P(x)={}^{n}C_{x}p^{x}q^{n-x}$, where $x = 0, 1, 2, 3 \ldots n$

P(x) is the probability mass function of the distribution. Mean of Binomial Distribution is np and variance is npq. Binomial Distribution is a discrete, probability distribution, which depends on only two parameters n and p.

16.6.2 : Poisson Distribution

Poisson Distribution is a limiting case of Binomial Distribution, when the number of trials n tends to become very large and probability of success (p becomes very small such that their product becomes a constant.

Probability that exactly x discrete events will take place is expressed by

$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ where λ = np (Constant), e is the base of the natural logarithms and $0 \leq x \leq \infty$

The mean of Poisson Distribution is and variance is also . It is interesting to see that Poisson Distribution has only one parameter and the value of mean and variance is the same.

16.6.3 : Normal Distribution

Normal distribution is symmetric with data more concentrated in the middle than in the tails. Normal Distribution is sometimes described as bell shaped. In order that a random variable is normal, it must satisfy the conditions like existence of a large number of chance factors, which are homogeneous and independent over the relevant population. The normal distribution is defined as follows:

If X is a random variable distributed normally with mean m and standard deviation σ, then its probability density function is given by

$$p(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \text{ where } -\infty < X < \infty,$$

Here π and e are constants with value 3.14 and 2.72 respectively. The distribution is completely known if the value of mean and standard deviation are known. The standard normal distribution is a normal distribution with a mean of 0 and a standard deviation of 1. Normal distributions can be transformed into standard normal distributions by the formula:

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

where X is a score from the original normal distribution, μ is the mean of the original normal distribution, and a is the standard deviation of original normal distribution. The standard normal distribution is sometimes also called the z distribution. The shape of the distribution is not be affected by this transformation. Areas under portions of the standard normal distribution are shown below:

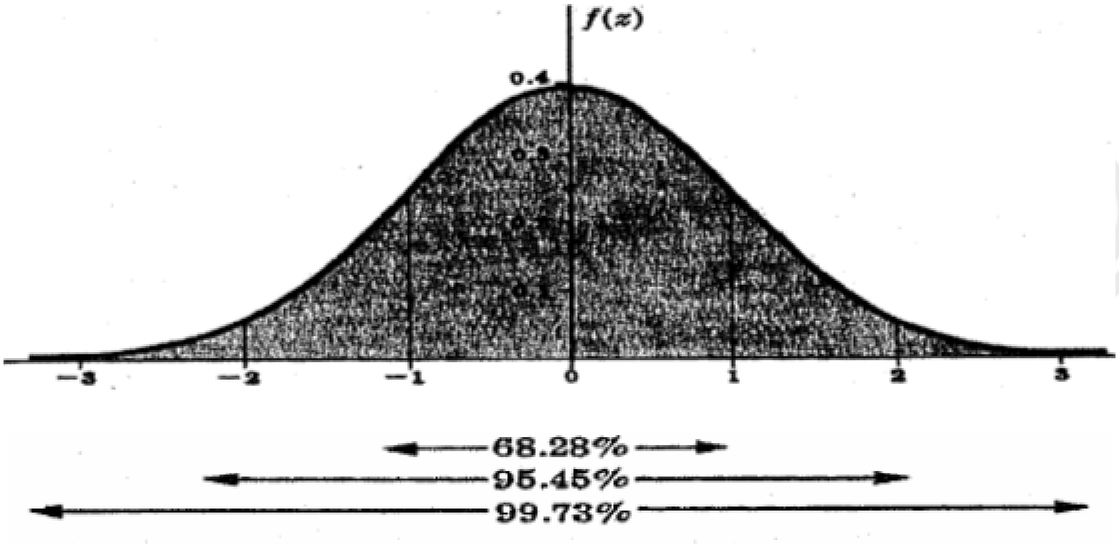

Fig. 1: Areas under portions of the standard normal distribution

The probability of a normal variate in an interval (XI, X_2) is given by the following formula

$$P(X_1 \leq X \leq X_2) = \int_{X_1}^{X_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} dX$$

This probability is equal to the area under the normal curve between the ordinates X_1 and X_2

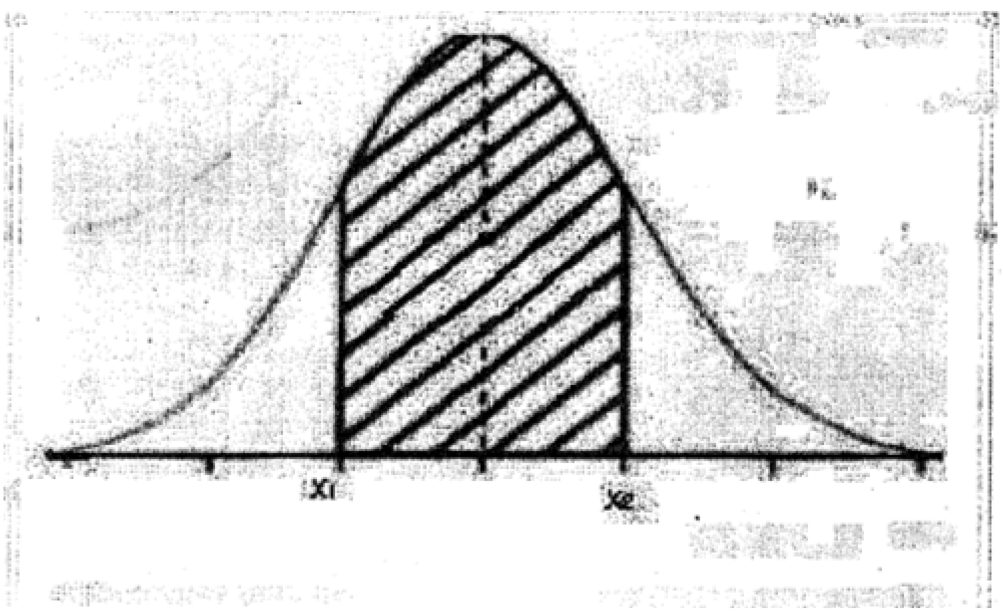

Fig. 2 : Area under the normal curve between the ordinates XI and X2

Alternatively we can write,

Here z is the standard normal variate and probability is equal to the area under the standard normal variate.

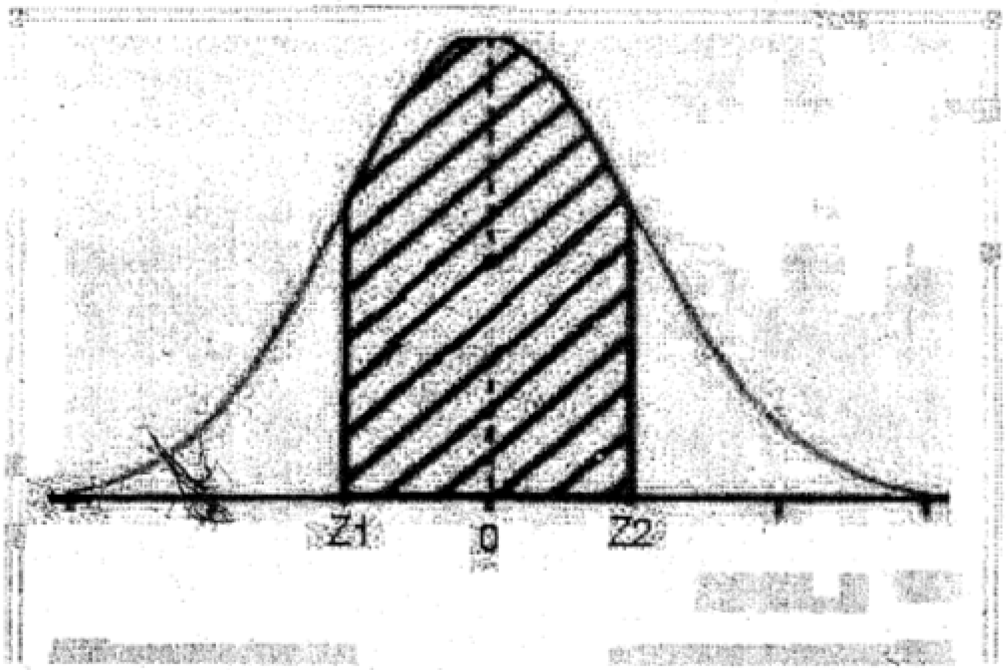

Fig. 3: Area under the standard normal variate

For example, to calculate the probability

$$P(-1 \le z \le 2) = P(-1 \le z \le 0) + P(0 \le z \le 2) = P(0 \le z \le 1) + P(0 \le z \le 2)$$

From 2 tables these probabilities are 0.3414 and 0.4772 respectively.

= 0.3414 + 0.4772 = 0.8185

16.7 : Applications of Probability

Probability is the mathematical theory, which is used to describe and quantify uncertainty. Uncertainty can be due to our lack of knowledge, deliberate amalgamation, or due to the essential randomness of "Nature". In any case, we measure the uncertainty of events on a scale from zero (impossible events) to one (certain events or no uncertainty).

Probability axioms form the basis for mathematical probability theory. Probability applications include even more than. Statistics, which is usually based on the idea of probability distributions. Probability theory plays a critical role in the development of statistical theory. Statistics is a branch of applied mathematics, which includes planning, summarizing, and interpreting uncertain observations. Whatever knowledge one has (about a process) is depicted mathematically and an attempt to learn more from whatever one can observe is made. This requires to:

a) plan observations to control their variability;
b) summarize a collection of observations to feature their community by suppressing details; and
c) reach consensus about what the observations tells about the world under observation.

In some forms of descriptive statistics, the data is often collected outside the control of the person doing the analysis, and the result of the analysis may be more an operational model than a consensus report about the world. Probability distributions are very useful in understanding the bibliometric "laws" that have been used to characterize counts of document-related preferences.

It helps in understanding the phenomenon used in explaining the bibliometric "laws". Probability distributions also have applications in modeling of behavior that identify other factors influencing the formation of individual preference. It forms the basis of testing of hypotheses for validating the model identified for a particular process. It has deep applications in webometrics, which is the application of traditional bibliometric techniques in analyses of the structure of the World Wide Web. In short, probability forms the basis for many Statistics methods, which are used in Library and Information Science.

16.8 : Summary

The theory of probability is a study of random experiment. An event with probability of 1 can be considered a certain event and an event with a probability of 0 can be considered an impossible event. In this Unit you have learnt about finding the probability of occurrence of an event. You have learnt about the various definitions, needed to understand the concept of probability. You also learnt about the additive and multiplicative property of the probability. You have understood what is meant by the probability distribution of a random variable and you have also studied about the joint probability and margins probability distribution. Finally you have studied about some special distributions, which are of immense importance in statistics. You would appreciate that a strong base in probability would help in understanding further issues like testing of hypothesis and decision theory better.