Block-I

Quantitative Methods

Introduction to Statistics and Probability

Block

I

INTRODUCTION TO STATISTICS AND PROBABILITY

Editorial Team

Prof. R. Prasad IFHE (Deemed-to-be-University), Hyderabad	Dr. V.S.S.N. Narasimha Murty Kadiyala IFHE (Deemed-to-be-University), Hyderabad
Prof. K. Seethapathi IFHE (Deemed-to-be-University), Hyderabad	Dr. Sanjay.Fuloria IFHE (Deemed-to-be-University), Hyderabad
Dr. Vishal Mishra IFHE (Deemed-to-be-University), Hyderabad	Prof. Muthu Kumar IFHE (Deemed-to-be-University), Hyderabad

Content Development Team

Dr. Kaustov Chakraborty IFHE (Deemed-to-be-University), Hyderabad	Prof. Muthu Kumar IFHE (Deemed-to-be-University), Hyderabad
Dr. Sashikala P IFHE (Deemed-to-be-University), Hyderabad	Dr. Y. V. Subrahmanyam IFHE (Deemed-to-be-University), Hyderabad

Proofreading, Language Editing and Layout Team

Ms.Jayashree Murthy IFHE (Deemed-to-be-University), Hyderabad	Mr. Venkateswarlu IFHE (Deemed-to-be-University), Hyderabad
Mr. Prasad Sistla IFHE (Deemed-to-be-University), Hyderabad	

Ref. No. QM SLM 102021B1

For any clarification regarding this book, the students may please write to The ICFAI Foundation for Higher Education (IFHE), Hyderabad specifying the unit and page number.

While every possible care has been taken in type-setting and printing this book, The ICFAI Foundation for Higher Education (IFHE), Hyderabad welcomes suggestions from students for improvement in future editions.

Our E-mail id: cwfeedback@icfaiuniversity.in

Center for Distance and Online Education (CDOE)
The ICFAI Foundation for Higher Education
(Deemed-to-be-University Under Section 3 of UGC Act, 1956)
Donthanapally, Shankarapalli Road, Hyderabad- 501203.

COURSE INTRODUCTION

Data based decision making is the current trend in the industry in 21st century. To analyze the data, statistical and mathematics techniques and tools are used. The management or business decision making uses many statistical techniques in areas such as finance, investments, operations, human resources, knowledge management, IT, and marketing. Statistical techniques can be used to analyze the data and identify the relationships between different variables in business domain. Useful conclusions and insights can be drawn using statistical techniques. In current days, computer based software applications are used for statistical programming and for efficient decision making in organizations.

Some of the concepts discussed in this course include arranging data, measures of central tendency and dispersion, probability, probability distribution and decision theory. Statistical inference, hypothesis testing using statistical techniques, correlation and linear regression are described in the course. Multiple regression, time-series analysis, quality control using statistics are also discussed in the course. Statistical techniques such as Chi-Square test, Analysis of Variance (ANOVA), index numbers, and linear programming are discussed. The role of IT in modern business enterprises, simulation and statistical software packages such as SPSS (Statistical Package for Social Sciences) and SAS (Statistical Analysis Systems) are explained in the course. Some of the advanced statistical techniques are also discussed.

Statistics play major role in management, business, economics, human resources, finance, marketing, applied psychology, and social sciences research. Statistics are widely used in business and management research and decision making. The business research process, research methodology, different types of research, survey methods, questionnaire design, statistical techniques used in data analysis are discussed in the course. Just doing the business or management research is not sufficient. One should be able to present the research findings in a report format. Different types of reports and their reporting styles are explained in the course.

This course consists of 6 blocks.

Block-1: Introduction to Statistics and Probability

Block-2: Statistical Relations and Hypothesis Testing

Block-3: Statistical Regression and Quality Control

Block-4: Statistical Distributions, Variations and IT

Block-5: Advanced Statistics

Block-6: Business Research

BLOCK I: INTRODUCTION TO STATISTICS AND PROBABILITY

The main aim of applying statistical techniques to the business is to study the relationship between the variables that affect the probability of the firm. Statistical tools are used to forecast future values of the variables that are included in financial decisions of the firm. To apply statistical tools and techniques, one should first collect the data and arrange it. Systematic classification and presentation of the collected data is very essential to process the data for decision making. This block deals with arranging data, central tendency and dispersion, probability, probability distribution and decision theory.

Unit-1 Arranging Data deals with the basic concepts of statistics and various tools used to arrange and present the collected data. In a business, statistics is used to study the business environment, to analyze business information and for many other purposes. For example, to analyze business revenues and to forecast sales, the demand and market characteristics of the product or service are to be analyzed. Statistics is a process of collecting, organizing and interpreting numerical facts. The collected data is arranged in a systematic way for analysis. Brilliant conclusions can be drawn, if the collected data is arranged in a systematic way.

Unit-2 Central Tendency and Dispersion discusses the importance of one of the most important objectives of statistical analysis, which is to get one single value that describes the characteristic of the entire mass of unwieldy data. Such a value is called the central value or an ‘average’ or the expected value of the variable. This single value is the point of location around which individual values cluster and, therefore, called the measure of location. Since this single value has a tendency to be somewhere at the center and within the range of all values, it is also known as the measure of central tendency. The very purpose of computing an average value for a set of observations is to obtain a single value which is representative of all the items.

Unit-3 Probability discusses the different types of events, different approaches to the probability, probability rules and Bayes’ Theorem. The chance of happening of an event is known as probability. By using probability concepts, the uncertainty in future can be eliminated.

Unit-4 Probability Distribution and Decision Theory discusses the random variables, probability distribution of a discrete random variable, discrete uniform distribution and binomial distribution. It also discusses Hypergeometric distribution, Poisson distribution, continuous uniform distribution, normal distribution, and lognormal distribution. t-distribution and F-distribution are discussed as well in the unit. The unit discusses the different features of random variables, and their application for decision-making. The variable in probability concepts is termed as the random variable. If a random variable can assume any value within a given range, it is called a continuous random variable. On the other hand, if the random variable can assume only a limited number of values, it is called a discrete random variable.

Unit 1

Arranging Data

Structure

1.1 Introduction

Statistics is a process of collecting, organizing and interpreting numerical facts. The collected data must be arranged in a systematic way to analyse the data. Brilliant conclusions can be drawn, if the collected data is arranged in a systematic way. Arranging the data is not so simple; therefore, so many tools and techniques are available to present the collected data for analysis and interpretation. In this unit, you will learn the basics of statistics and various tools to arrange and present the collected data.

1.2 Objectives

After going through the unit, you should be able to:

- Define statistics;
- State differences between quantitative and qualitative data;
- State the need for arranging data;
- Explain Frequency distribution;
- Formulate histogram;
- Label frequency polygons; and
- Label ogive curves.

1.3 Meaning and Definition of Statistics

The Webster's Dictionary defines 'Statistics' as "The science which has to do with the collection, classification and analysis of facts of a numerical nature regarding any topic." In general, statistics is divided into two parts – descriptive statistics and inferential statistics.

'Descriptive Statistics' summarizes the data with a sample information that characterizes the whole data. It also refers to the presentation of a body of data in the form of tables, charts, graphs, etc. 'Inferential Statistics' refers to drawing generalizations about the properties of the whole population from a sample drawn from the population. It is the process of making estimation, prediction or decision about the population. Let us familiarize ourselves with some important statistical terms.

Population: Population is complete collection of all elements in a statistical problem; for example, all students of an educational institution, the salaries of all workers in a community, etc.

Census: A census is the process of collection of data from every element in a population.

Sample: Sample is a sub-collection of data from the population. For example, the salary of all workers in a community is population but collection of salaries of a few workers for estimation about the whole population is called a sample. Sample is always a subset of the entire population.

Parameter: Parameter is a numerical measurement that describes the characteristics of a population.

Statistic: Statistic is a numerical measurement that describes the characteristics of a sample.

The following Exhibit 1.1 will facilitate us to understand the basic terms in statistics.

Exhibit 1.1: Post-poll Surveys by CSDS

Post-poll surveys were conducted by many research institutions. One such study was by the Centre for the Study of Developing Societies (CSDS), an Indian research institute through its research programme Lokniti along with The Hindu, in 4 different states - Assam, Kerala, Tamil Nadu and West Bengal in 2021 where the Assembly elections were conducted. The surveys were conducted from March 28, 2021 through May 1, 2021 in different phases in the states in the local languages. The survey was conducted with 3473 voters in Assam, 3424 voters in Kerala, 4354 voters in Tamil Nadu and 4223 voters in West Bengal.

Contd….

The study report presented many tables and charts based on the responses to various questions. The study concluded that the people verdicts in these states were based on the respective local issues in each state.

Source: https://www.thehindu.com/opinion/op-ed/local-factors-determine-electoral-outcomes-in-states/article34475075.ece and https://www.lokniti.org/POST_POLL_ANALYSIS_2021

This exhibit helps us to familiarize with a few important terms in statistics as given below:

As given in the exhibit, sample data was collected by CSDS, then the data was classified and analysed. As the data was presented in many tables and charts, it can said that the study used descriptive statistics.

Population: All eligible voters.

Population parameter: The proportion of all eligible voters who favor a particular candidate.

Sample: The eligible voters surveyed - 3473 voters (Assam), 3424 voters (Kerala), 4354 voters (Tamil Nadu) and 4223 (West Bengal).

Sample statistics: The proportion of eligible surveyed voters in each state who favor a particular candidate's win.

Quantitative and Qualitative Data: Characteristics which cannot be expressed numerically are called 'qualitative data' and those which can be expressed numerically are called 'quantitative data'.

1.3.1 Statistical Application in Business

Statistics can be used in business in various ways and areas. These areas may be financial management, marketing management, operational management, etc. In financial management, statistics can be used for capital budgeting, capital structuring, working capital management, stock and bond valuation, capital market etc.

The capital budgeting process often requires to forecast the future revenue and operation for judging the feasibility of the project. Probability is used for this purpose. For capital structuring process, we often require to value the bonds and stocks based on their future or past statistics. Similarly, we use statistics in working capital management to choose optimal short-term financing. This can be used to reveal long-term trends and seasonal variations in sales, expenses and incomes. Valuation of stocks and bonds also require statistical data. Besides financial management, statistics has several applications in the area of marketing management, operational management, human resources management, etc.

1.4 Frequency Distribution

A frequency distribution shows the number of observations falling into each of several ranges of data points. This arrangement is done through preparation of a frequency distribution table. A frequency distribution table is a way of organizing the data by listing every possible data points as a column of numbers and the frequency of occurrence of each data point as another. Computing the frequency of a score is simply a matter of counting the number of times that score appears in the set of data. It is necessary to include scores with zero frequency in order to draw the frequency polygons correctly.

1.4.1 Arranging Data

A financial analyst collected the sales data of a large engineering company for 20 days. These collected data sometimes were large and filled a number of pages and did not make sense to the analyst. Statistics uses a technique that helps to condense the data using tables, graphs, etc. The data which is originally collected is called 'raw data'.

Example 1

Raw data of sales (Rs. in lakh) of a large engineering company for the 20 days (April 20xx) is given below:

14 8 23 31 26 5 11 29 46 32

28 12 36 8 9 16 42 30 9 22

In the above example, one cannot say for how many days the sales of companies is between 20 and 30 lakh rupees, or what the minimum and maximum sales during the period, etc., are Frequency distribution exhibits how frequencies are distributed over various categories. It represents the results in a more simple and meaningful way. The above presentation of data can be improved by creating an array in which the prices are arranged in ascending or descending order.

The frequency distribution of the above example is given below:

Sales (Rs. in lakh)	Tally Mark	Frequency (f)
1-10	𝍸	5
11-20	||||	4
21-30	𝍸 |	6
31-40	|||	3
41-50	||	2
Total		20

For every raw data, we mark a tally in the second column so this is called tally mark. Frequency is the total number of tally in a given interval and is denoted by f. The difference in the higher and lower value of the first column is called as ‘class interval’.

The steps given below can be followed for preparing a frequency table:

1. Specify the number of class intervals. There is no specific rule that tells us how many intervals are to be used. But the number of classes should usually be between 6 and 15. The classes must be both mutually exclusive and all-inclusive. Mutually exclusive means that the classes must be selected such that an item cannot fall into two classes and all-inclusive classes are classes that together contain all the data.
2. When all intervals are to be of the same width, the following rule may be used to find the required class interval width:

   W = (L – S)/N

   Where,

   W = class width,

   L = the largest data,

   S = the smallest data,

   N = number of classes.

   Sometimes situations do arise when a class interval has to be kept open-ended. The open class may be the first class or the last. When first class is open we can write less than when the last class is open; we can write more than for the latter. An open interval is needed when the few observations in a given data are either too small or too large. For example, if we have to make a frequency distribution of Rs.12,000, Rs.20,000, Rs.15,000, Rs.21,000, Rs.2,000 and Rs.60,000. We take first interval as less than 5,000 and last interval as Rs.20,000 and above.

Following points should be remembered while preparing frequency distribution:

- Every item of data or data point should be included in one, and only one class and the lowest share price should be included in the first class and the highest share price in the last class.
- Types of classes where the upper limit of one class equals the lower limit of the next class are called “exclusive classes” because the upper limit of a class is excluded from the class.
- Types of classes where, the upper limit of each class is included in that class is called “inclusive class”.

- Class intervals should be of the same length to the extent possible.
- Number of classes can be the square root of total number of frequencies.

Check Your Progress - 1

1. Which of the following is not within the scope of Statistics?

   a. Collecting data

   b. Interpreting data

   c. Selecting tools

   d. Organizing numerical facts

   e, Classifying data

2. Fill up the blank with a suitable term.

   Summarization of data with a sample information that characterizes the whole data comes under….

3. Characteristics which cannot be expressed numerically is called……..

4. Which activity from the following cannot be used for applying statistics to financial management?

   a. Capital budgeting

   b. Capital structuring

   c. Working Capital Management

   d. Capital market

   e. Settling suppliers' dues

5. What is the drawing of generalizations about the properties of the whole population from its sample is referred as

   a. Descriptive statistics

   b. Inferential statistics

   c. Quantitative statistics

   d. Qualitative statistics

1.5 Graphical Frequency Distribution

As we discussed earlier, data may be represented using a table or graph. Tabular representation of data sometimes may not provide the important information that graphical representation can give. So, graphical representation of data is a powerful tool that helps in various decisions-making situations. Some of the frequently used methods to present the data in graphical forms are as follows:

1.5.1 Histogram

A histogram is a way of summarizing data that are measured on an interval scale (either discrete or continuous). But histogram is normally used for continuous variables. We can graph this frequency distribution by taking classes or class marks (mid-points of classes) on the X-axis and frequencies on the Y-axis.

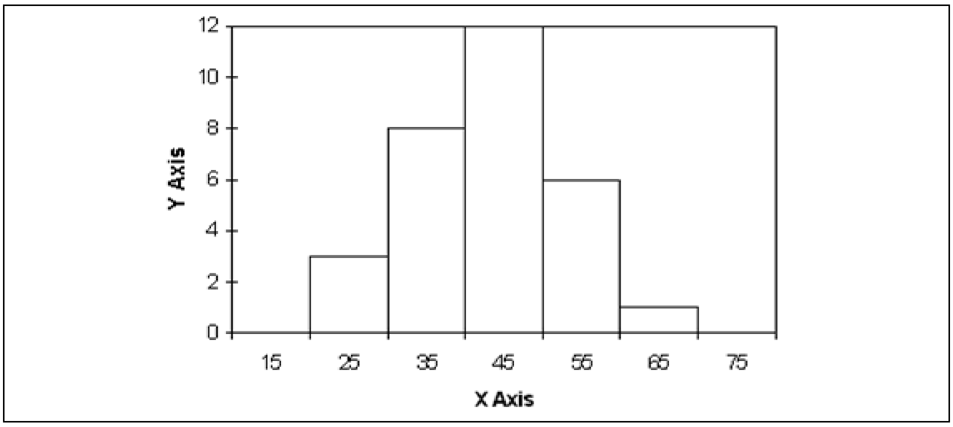
Consider the following frequency distribution table:

Table 1.1

Class Interval	Frequency
20-30	3
30-40	8
40-50	12
50-60	6
60-70	1

The histogram for the given data from the above frequency distribution table is shown below.

Figure 1.1: Histogram



Following are the characteristics of histogram: (i) Classes are represented by the base of the rectangles and frequencies are represented by the heights of the rectangles. (ii) If the classes of histogram are of equal width then bases of the rectangles will be of equal length too. (iii) The tallest rectangle in a histogram represents the class with the highest frequency.

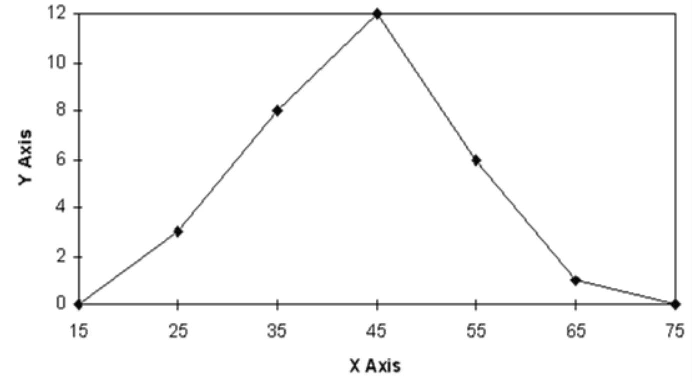
Demerits of Histogram

While the histogram indicates the fluctuations in frequencies from class to class, it does not clearly reveal the rate of change in frequency from one class to the next.

1.5.2 Frequency Polygons

We can graph the frequency distribution by taking classes or class marks (mid-points of classes) on the X-axis and frequencies on the Y-axis. The frequency polygon for the frequency distribution given in figure 1.2 is shown below.

Figure 1.2: Frequency Polygon



1.5.3 Ogive Curve

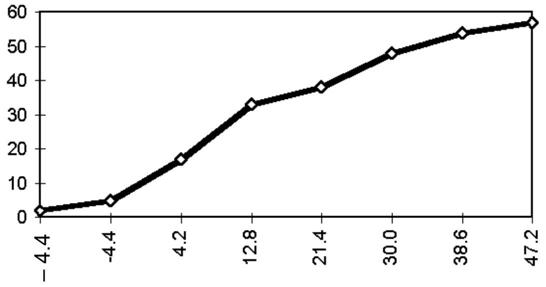
Cumulative frequencies are plotted at the class marks and successive points are connected by straight lines. Cumulative frequency may be of less than or more than type, while ogive can also be of more than type or less than type. Consider the following cumulative frequency distribution:

Table 1.2

Class	Cumulative Frequency
–13.2 ≤ x < –4.4	2
–4.4 ≤ x < –4.2	5
– 4.2 ≤ x < 12.8	19
–12.8 ≤ x < 21.4	33
–21.4 ≤ x < 30.0	38
–30.0 ≤ x < 38.6	48
–38.6 ≤ x < 49.2	54
–49.2 ≤ x < 55.8	59

Ogive of the cumulative frequency given above in table 1.3 is shown below:

Figure 1.3: Ogive Curve



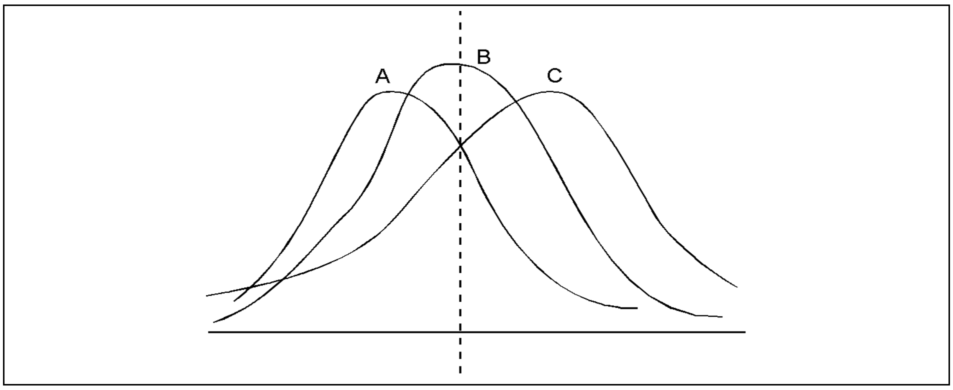
It may be noticed from the "less than" ogive curve below that it slopes up to the right.

1.6 Skewness

Skewness refers to the lack of symmetry. A distribution for which the mean, median and mode are equal is known as a 'symmetrical distribution'. In such a distribution curve, a vertical line drawn from the peak of the curve to the horizontal axis will divide the area of the curve into two equal parts and each part is the mirror image of the other. An asymmetrical distribution for which the mean, median and mode are not equal is known as a 'skewed distribution'. In a skewed distribution curve, the values are not equally distributed but are concentrated at the lower- or higher-end of the frequency distribution.

In at the lower-end and very few values are at the higher-end, the curve is said to be skewed to the right or positively skewed. For a negatively skewed curve, the values are concentrated at the higher-end and it is skewed to the left because it tails off towards the lower-end. B is symmetrical while A is said to be skewed to the right and C is skewed to the left.

Figure 1.4: Symmetrical & Asymmetrical Distributions



<u>Check Your Progress - 2</u>

6. If a frequency distribution curve tails off to the right, then it is said to be
   a. Symmetrical
   b. Skewed to the right
   c. Skewed to the left
   d. Negatively skewed
   e. None of the above.
7. The skewness of a distribution is indicated by
   a. Histogram
   b. Ogive
   c. Frequency polygon
   d. Cumulative frequency table
   e. Cumulative frequency curve.

8. For a grouped data, the graph plotted by taking mid-points of each class on X-axis and frequency on Y-axis is called
   a. Histogram
   b. Frequency polygon
   c. Frequency curve
   d. Frequency chart
   e. Frequency table.
9. An Ogive is
   a. A graph of ungrouped data
   b. A graph of grouped data
   c. A graph of cumulative frequencies
   d. A graph of ranges of fractiles
   e. A graph with rectangles as opposed to a line graph.
10. If a frequency distribution has more values at the right end, it is called
   a. Symmetrical distribution
   b. Positively skewed
   c. Negatively skewed
   d. Skewed to the right
   e. None of the above

1.7 Summary

- Statistics is divided into two basic areas – descriptive statistics and inferential statistics. Descriptive statistics involves arranging, summarizing and presenting a set of data in such a way that a meaningful data can be produced and interpreted. Inferential statistics refers to the drawing of generalizations about the properties of the whole population from its sample.
- Statistical data may be represented either through a table or a graph. Frequency polygon, histogram and ogive are some of the frequently used graphs and are the most powerful tools in decision-making.

1.8 Glossary

Class: A class refers to a group of objects with same common property.

Class Limit: The range of values of a given class is called a ‘class limit’, and middle of a class interval is called ‘class mark’.

Cumulative Frequency Distribution: A cumulative frequency distribution is a tabular display of data showing how many observations lie above or below, certain values.

Discrete Data: Discrete data takes value only in a whole number.

Frequency Distribution: In frequency distribution table, raw data is tabulated by dividing it into classes of convenient size and computing the number of data elements falling within each pair of class boundary.

Frequency Polygon: It is a graph of frequency that connects the mid-point of each data set, plotted at a height corresponding to the frequency of the class.

Histogram: A histogram is a way of summarizing data in the form of rectangle measured on an interval scale.

Ogive: Ogive is a graph of cumulative frequency.

1.9 Suggested Readings/Reference Material

1. Gupta, S. P. Statistical Methods. 46th Revised ed. New Delhi: Sultan Chand & Sons. 2021.
2. I. Levin Richard, H. Siddiqui Masood, S. Rubin David, Rastogi Sanjay. Statistics for Management. Pearson Education; Eighth edition, 2017.
3. Gerald Keller. Statistics for Management and Economics. Cengage, 2017.
4. Arora, P. N., and Arora, S. CA Foundation Course Statistics. 6th ed. S Chand Publishing, 2018.
5. Mario F Triola. Elementary Statistics. 13th ed., 2018.
6. David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran. Statistics for Business and Economics. 13th Edition, Cengage Learning India Pvt. Ltd., 2019.
7. S D Sharma. Operations Research. Kedar Nath Ram Nath, 2018.
8. Hamdy A. Taha. Operations Research: An Introduction. 10th ed., Pearson, 2016.
9. Malhotra, N. (2012), Marketing Research: An Applied Orientation, 7th ed., Pearson, 2019.
10. Cooper, D.R. and Schindler, P.S. and J. K. Sharma (2018), Business Research Methods, 12th edition, McGraw-Hill Education.

1.10 Self-Assessment Questions

1. The Managing Director of a dairy company is planning to start a dairy product. His marketing team has made a random survey of 2000 offices in Andhra Pradesh to check the number of packets of milk they buy. The survey report is as follows:

Number of Milk packets	1-2	2-3	3-4	4-5	5-6
Number of outlets	540	820	310	210	120

Draw the histogram and ogive from these data.

2. Explain the differences between descriptive statistics and inferential statistics.

3. What do you mean by exclusive class and inclusive class?

4. Explain the term “skewness”.

5. Describe the properties of a histogram.

1.11 Answers to Check Your Progress Questions

1. (c) Selecting tools

2. Descriptive statistics

3. Qualitative data

4. (e) Settling suppliers dues

5. (b) Inferential statistics

6. (b) Skewed to the right

7. (c) Frequency polygon

8. (b) Frequency polygon

9. (c) A graph of cumulative frequencies

10. (c) Negatively skewed

Unit 2

Central Tendency and Dispersion

Structure

2.1 Introduction

In the previous unit, you learnt basis terms of statistics and the arrangement of data using frequency polygon, histogram, etc. The main objective of this unit is to understand the importance of central tendency and dispersion. The measure of central tendency is a single value that is used to represent whole data set. This unit exclusively covers of the use the statistical measures in the area of finance.

2.2 Objectives

After going through the unit, you should be able to:

- Evaluate the need for measuring central tendency;
- Recall the concepts of Mean, median and mode and other statistical central tendency techniques;
- Define range and standard deviation;
- Define the Bienayme-Chebyshev Rule; and
- Explain coefficient of variation.

2.3 Central Tendency

After calculating frequency distribution, the next task of a statistician is to compress the data further into one component to represent the characteristics of the entire data. Measures of central tendency are measures that provide one representative value for the data set. Since an average represents the entire data, its value lies somewhere in between the two extremes, i.e., the largest and the smallest items. For this reason, an average is frequently referred to as a measure of central tendency.

Objectives of Measuring Central Tendency

As discussed earlier, the objective of measuring central tendency is to arrive at a single value that represents the whole data. The main objectives of averaging include:

i. **To Find a Single Value that Represents the Whole Data:** It is very difficult to analyze if the number of data is high. For example, it is very difficult to remember the sales of all companies in an industry. By averaging a single number can be used to represent the sales of an industry.

ii. **To Facilitate Comparison:** Average is used as a common measure to compare two or more sets of data. Further, it can be helpful in taking decisions about the characteristics of different sets of data. For example, a Sales Manager can compare sales of different areas to compare the performance of the sales executive of the respective areas.

Requisites of a Good Central Tendency

A central tendency should be: (a) vigorously defined, (b) easy to compute, (c) capable of simple interpretation, (d) dependent on all the observed values, (e) not unduly influenced by one or two extremely large or small values, (f) fluctuate relatively less from one random sample to another, and (g) capable of mathematical manipulation.

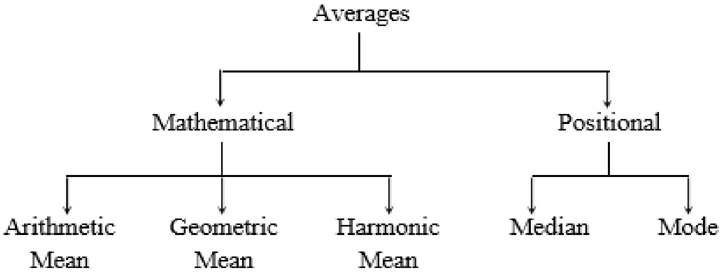
Types of Averages

The following are the various types of common means used in the statistical analysis:

Mathematical: Arithmetic Mean, Geometric Mean and Harmonic Mean

Positional: Median and Mode

Figure 2.1: Flowchart of Averages



2.4 Arithmetic Mean

The other familiar term for arithmetic mean is average. For n observations X_1, X_2, X_3 …..X_n, the arithmetic mean is computed by summing the observations and dividing by the number of observations. Mathematically, it can be represented by the formula:

$$\overline{X} = (X_1 + X_2 + ...+ X_n)/n = \left(\sum_{i=1}^{n} X_i\right)/n$$

Where,

n is the number of observations and

the variable X takes the values X_1, X_2, ... X_n.

Mean of population is represented by μ and the mean of sample is represented by $\overline{X}$

Example 1

The following table gives the profit of ABC Ltd. for the period of 10 years.

Year	1	2	3	4	5	6	7	8	9	10	Total
Profit (Rs. in crore)	425	432	444	450	452	455	459	465	470	475	4,527

Now, the arithmetic mean of profit of ABC Ltd., for the period of 10 years can be calculated as follows:

$$\text{Arithmetic Mean} = \frac{\text{Sum of all the profits}}{\text{No. of years}} = \frac{4,527}{10} = 452.7 \text{ crore.}$$

2.4.1 Calculation for Discrete Series or Ungrouped Data

For discrete series or ungrouped data, the mean can be calculated by using the given formula:

$$\overline{X} = \frac{\sum fX}{\sum f}$$

Where,

f = frequency and X = variable.

Example 2

In a survey of 50 steel companies, the following data was collected about the level of profits attained by them:

X_i = Level of profit (Rs. lakh) earned during 20x1-20x2	f_i = No. of companies that earned X_i amount of profit	X_if_i
25	20	500
15	16	240
22	15	330
21	14	294
35	16	560
Total	81	1,924

The arithmetic mean is:

$$\overline{X} = \frac{\sum fx}{\sum f} = \frac{1,924}{81} = 23.75$$

Thus, the average profit of steel industries is Rs.23.75 lakh.

2.4.2 Calculation for Continuous Series or Grouped Data

In case of continuous series or grouped data, mean can be calculated as:

$$\overline{X} = \sum fm/N$$

Where,

m = mid-point of class = $\frac{\text{Lower limit} + \text{Lower limit of next class}}{2}$

f = frequency of each class

N = Total frequency = $\sum f$.

Example 3

The Managing Director of a reputed newspaper is planning to start a new business daily. His marketing team has made a random survey of 2,000 offices in Maharashtra to check the number of daily newspapers they buy. The survey report is as follows:

Number of papers	1-2	2-3	3-4	4-5	5-6
Number of offices	540	820	310	210	120

Computation of Mid Value:

Class	Mid-point (m)	Number of offices (f)	fm
1-2	1.5	540	810
2-3	2.5	820	2,050
3-4	3.5	310	1,085
4-5	4.5	210	945
5-6	5.5	120	660
Total		2,000	5,550

$$\overline{X} = \frac{\sum fm}{\sum f} = \frac{5,550}{2,000} = 2.775$$

Where, $\sum f = N$ = Total no. of Offices. Thus, the mean is 2.775.

2.4.3 Properties of Mean

i. The sum of deviations of the items from the arithmetic mean (taking signs into account) is always zero, i.e., $\sum (X - \overline{X}) = 0$.

ii. The sum of the squared deviations of the items from arithmetic mean is less than the sum of the squared deviations of the items from any other value, i.e., $\sum (X - \overline{X})^2$ is less than $\sum (X - A)^2$ where A is any other point, different from $\overline{X}$.

iii. Since, $\overline{X} = \sum X/N$, $(N \text{ x } \overline{X}) = \sum X$.

iv. If we have the arithmetic mean and number of items of two or more than two groups, we can compute the combined average of these groups, by applying the following formula:

$$\overline{X}_{1,2} = (N_1 \overline{X}_1 + N_2 \overline{X}_2)/(N_1 + N_2)$$

Where,

$\overline{X}_{1,2}$	=	Combined mean of the two groups,
$\overline{X}_1$	=	Arithmetic mean of the first group,
$\overline{X}_2$	=	Arithmetic mean of the second group,
N_1	=	No. of items of the first group,
N_2	=	No. of items of the second group.

Uses

Arithmetic mean is widely used because of the following reasons: (i) It is easy to compute and to understand and is rigidly defined so that different interpretations by different persons are not possible. (ii) It is relatively reliable as it does not vary too much as some other statistical descriptions when repeated. Samples can be taken from one and the same population. (iii) It is typical in the sense that it is the center of gravity balancing the values on either side of it. (iv) It takes all values into account, so it is more meaningful.

Abuses

Calculation of Mean may be simple and foolproof, but the application of result may not be so, foolproof. It can often be misleading if the data do not fall in a homogenous group.

2.4.4 Weighted Arithmetic Mean

Weighted arithmetic mean is taken into consideration when relative weight is assigned to each of the values. Instead of assigning equal weight in arithmetic mean weighted arithmetic mean assigns different values. The formula for computing weighted arithmetic mean is:

$$\overline{X}_W = \frac{\sum WX}{\sum W}$$

Where,

$\overline{X}_w$ represents the weighted arithmetic mean;

X represents the variable values, i.e., X_1, X_2...X_n;

W represents the weight assigned to the value.

Weighted average mean is used to calculate the index numbers. Comparison based on weighted average mean is more meaningful.

2.5 Median

Median is a positional measure that divides the distribution into two equal parts (in a number of cases) when the data can be arranged in a rank order from low to high. In other words, 50% of the sample falls below the median and 50% of the sample falls above the median and median is just the 50th value of the series. Consider the five numbers 1, 2, 3, 4, 5. 3 is the median, as well as the mean but if you consider the five numbers 1, 2, 3, 7, 12, the median is still 3 but mean is 5.

2.5.1 Ungrouped Data

The first step in the calculation of median of ungrouped data is to arrange the series in an ascending order. When the totals of the list are odd, the median is the middle entry in the list. When the totals of the list are even, the median is equal to the average of the two middle entries.

If the total of the frequencies is odd, say n,

Median = $[(n + 1)/2]^{th}$ item

If the total of the frequencies is even, say, 2n, then,

Median = Average of n^{th} and $(n + 1)^{th}$ item.

2.5.2 Grouped Data

Median of the grouped data can be calculated by using the following formula:

$$\text{Median} = \left[\frac{(N+1)/2 - (F+1)}{f_m}\right] w + L_m$$

Where,

L_m	=	Lower limit of the median class
f_m	=	Frequency of the median class
F	=	Cumulative frequency up to the lower limit of the median class
W	=	Width of the class interval
N	=	Total frequency.

Median is the size of the $\left(\frac{N+1}{2}\right)^{th}$ item.

Example 4

The net profit margin of 100 companies is given below.

Net Profit Margin (%)	0-10	10-20	20-30	30-40	40-50
No. of Companies	12	18	25	20	25

Find the median net profit for the given data.

Net Profit Margin (%)	No. of Companies (f)	Cumulative Frequency (cf)
0-10	12	12
10-20	18	30
20-30	25	55
30-40	20	75
40-50	25	100

Here, the total frequency N = 100. Median is the size of the (N + 1)/2th item, i.e., [(100 + 1)/2]th item, i.e., the size of the 50.5th item. It lies in the class 20-30. Hence, 20-30 is the median class, of which the lower limit is 20.

Thus, L_m = 20, N = 100, F = 30, f_m = 25, W = 10

Substituting these values in the above formula

$$\text{Median} = \left[\frac{(N+1)/2 - (F+1)}{f_m}\right] W + L_m = \left[\frac{(100+1)/2 - (30+1)}{25}\right] 10 + 20 = 27.8$$

Thus, 27.8 is the median net profit margin of the companies.

Properties of Median

The sum of deviations of the items from the median (ignoring signs) is always minimum i.e., zero.

Advantages

Median has several advantages over the others: (i) Median is easy to understand and calculate as it indicates the value of the middle item in the distribution. (ii) Median is not influenced by the extreme value as strongly as the mean. So, it is an appropriate measure for income distributions or price distributions where the arithmetic mean would be distorted by extreme values. (iii) Median can be calculated from any type of data, even for grouped data with an open end interval, unless the median falls into an open-ended class. (iv) Median is the most appropriate average in dealing with qualitative data, i.e., where ranks are given or when items are not counted or measured, but scored. (v) The value of median can be determined graphically, whereas the value of mean cannot be graphically ascertained.

Disadvantages

i. Median can be erratic and give in accurate values when the number of items is small. In addition, since median is calculated for even and odd numbers in different ways, it has been less historically used than mean as a basis of building statistical interference.

ii. Median is a positional average so it is not determined by taking every observations into account.

iii. Arrangement of data either in ascending order or descending order is necessary for calculating the median which is a time consuming process. It is very difficult to do so if the number of observations are large.

iv. Median is not suitable for algebraic treatment. For example, combined median of two or more than two groups, is not possible. Because of this limitation, median is much less popular as compared to the arithmetic mean.

v. The value of median is affected more by sampling fluctuations than the value of arithmetic mean.

2.6 Related Positional Measures

In addition to the above, some other positional measures like quartile, decile, percentile are also in use. The methods of computation of these positional measures are given below. Quartiles, deciles, etc., are computed just Q as the median. While computing these values in individual and discrete series we add 1 to N whereas in continuous series we do not add 1. Thus,

Q_1 = Size of $\left(\frac{N+1}{4}\right)^{th\ th}$ item (individual observation and discrete series)

Q_3 = Size of $\left(\frac{3(N+1)}{4}\right)^{th\ th}$ item (in individual and discrete series)

th

Q_3 = Size of $\left(\frac{3N}{4}\right)$ item (in continuous series)

D_4 = Size of $\left(\frac{4(N+1)}{10}\right)^{th}$ item (in individual and discrete series)

D_5 = Size of $\left(\frac{4N}{10}\right)^{th}$ item (in continuous series)

P_{60} = Size of $\left(\frac{60(N+1)}{100}\right)^{th}$ item (in individual and discrete series)

P_{60} = Size of $\left(\frac{60N}{100}\right)^{th}$ item (in continuous series).

Example 5

Calculate the lower and upper quartiles, third deciles and 20th percentile from the following data:

Central value	2.5	7.5	12.5	17.5	22.5
Frequency	14	36	50	60	40

Since mid-points are given, we have to find the lower and upper limits of the various classes. Take the difference between the two central values and divide it by 2, then deduct the value so obtained from the lower limit and add it to the upper limit. In the given case:

$$\frac{7.5-2.5}{2} = \frac{5}{2} = 2.5$$

The first class shall be 0-5 and second 5-10, etc.

Calculation of Q_1, Q_2, D_3, P_{20}

Class group	f	c.f.
0-5	14	14
5-10	36	50
10-15	50	100
15-20	60	160
20-25	40	200
	N = 200	

Lower Quartile Q_1 = Size of $\left(\frac{I^{th}}{4}\right)$ item = $\frac{200}{4}$ = 50th item

Q_1 lies in the class 5-10

$$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} x\,i$$

L = 5, $\frac{N}{4}$ = 50, c.f. = 14, f = 36, i = 5

$$Q_1 = 5 + \frac{50-14}{36} x 5 = 10$$

Upper Quartile Q_3 = Size of $\left(\frac{3]^{th}}{4}\right)$ item = $\left(\frac{3\,x\,2(^{th})}{4}\right)$ item = 150th item

Q_3 lies in the class 15-20.

$$Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} x i$$

L = 15, $\frac{3N}{4}$ = 150, c.f. = 100, f = 60, i = 5

$$Q_3 = 15 + \frac{150-100}{60} x 5 = 19.17$$

Third Decile D_3 = Size of $\left(\frac{3\Gamma^{th}}{10}\right)$ item = $\frac{3 \times 200}{10}$ = 60th item

D_3 lies in the class 10-15,

$$D_3 = L + \frac{\frac{3N}{10} - c.f.}{f} \times i$$

L = 10, $\frac{3N}{10}$ = 60, c.f. = 50, f = 50, i = 5

$$D_3 = 10 + \frac{60-50}{50} \times 5 = 11$$

Twentieth Percentile P_{20} = Size of $\left(\frac{20^{th}}{100}\right)$ item = $\frac{20 \times 200}{100}$ = 40th item

$$P_{20} = L + \frac{\frac{20N}{100} - c.f.}{f} \times i$$

L = 5, $\frac{20N}{100}$ = 40, c.f = 14, f =36, i = 5

$$P_{20} = 5 + \frac{40-14}{36} \times 5 = 8.61.$$

2.7 Mode

Mode is a value that appears most frequently in the distribution. The calculation of mode differs from the mean as it does not require a basic arithmetical process. For example, mode of the observation 1, 2, 2, 3, 4, 5, 4, 6, 4, is 4 as 4 is repeated 3 times i.e., more than any other number.

Example 6

The following data relates to the intraday share price quotations of Reliance Industries Ltd. quoted on 17th September, 2021.

2428, 2430, 2435, 2450, 2450, 2441, 2450, 2445, 2445, 2444, 2440, 2439, 2438, 2437, 2437, 2450.

Let us calculate the mode for the above data.

Price quotation (Rs.)	No. of times it occurs
2428	1
2430	1
2435	1
2437	2
2438	1
2439	1
2440	1
2441	1
2444	1
2445	2
2450	4
Total	16

Since the share price 2450 has occurred the maximum number of times, the mode is 2450.

2.7.1 Grouped Data

For grouped data the first task is to identify the modal class, which is where the most number of observations occur. For calculating mode from a frequency distribution, the following formula can be used:

$$\text{Mode} = L_{mo} + \frac{f - f_1}{2f - f_1 - f_2} \text{ x } W$$

Where,

L_{mo} = Lower limit of the modal class which is the class having the maximum frequency.

f_1, f_2 = Frequencies of the classes preceding and succeeding the modal class respectively.

f = Frequency of the modal class.

W = Class interval.

Example 7

Profit of 50 companies that belong to Nifty is given below:

Profit in crore	0-10	10-20	20-30	30-40	40-50
No. of companies	12	16	11	4	6

Let us find the mode for the above data: Here the largest frequency is 16, and lies in the class 10-20. So the modal class is 10-20. Therefore,

$L_{mo} = 24, f_2 = 28, f_1 = 36, W = 8, f = 43$

$$\text{Mode} = L_{mo} + \frac{f - f_1}{2f - f_1 - f_2} \text{ x } W$$

$$= 10 + \frac{16-12}{2 \text{x} 16 - 12 - 11} \text{ x } 10 = 10 + 4.4 = 14.44.$$

2.7.2 Bimodal Distribution

Sometimes, it is possible to have two frequencies for the same class and that may also be the highest. The distribution is then called bimodal because of two modes. A graph of such distribution will have two highest points. Here, the value of the mode cannot be determined with the help of the formula given above. It is advisable to use other measures of central tendency in case of a bimodal distribution.

Advantages

i. Mode is most useful in qualitative measurement. For example, if a Sales Manager wants to survey the preference for a product mode is a better measure than others.

ii. Like the median, mode is not affected by the extreme value. Be they large or small values too do not matter. For example, mode of 1, 1, 2, 2, 2, 10 is 2 and mode of 12, 15, 12, 12, 10 is also 12.

iii. Like the median, mode can also be determined for an open-ended interval.

iv. The value of mode can be determined graphically, whereas the value of mean cannot be graphically ascertained.

Disadvantages

i. In case of bimodal or multimodal distribution, mode cannot be determined.

ii. We cannot calculate the combined mode for two groups, because mode is not capable of algebraic manipulations.

iii. The value of mode is not based on each and every item of the series.

2.7.3 Empirical Mode

The value of mode may be defined by ascertaining the empirical relation governing Mean, Median and Mode: Mode = 3 x Median – 2 x Mean. This measure is called the empirical mode.

2.8 Appropriateness of the Three Principal Averages

Sometimes it is difficult to choose which measure to use. In fact mean, median and mode have their own advantages. The proper use of central tendency depends on the situation.

The mean is ordinarily the preferred measure of central tendency and is the arithmetic average of a distribution. The mean, presented along with the variance and the standard deviation, is the "best" measure of central tendency for continuous data. But in some situations, median or mode is preferred to a mean. Such situations are: (i) when distribution is skewed. (ii) when the number of objects is very small. (iii) when the interval is open-ended.

The mode is not frequently used as a measure of central tendency because the largest frequency of scores might not be at the center. But, in the situations where describing discrete categorical data as the greatest frequency of responses is important, mode is the best measure.

2.8.1 Comparison of the Principal Averages

- The mean, median and mode are located at the same point in a symmetrical frequency distribution.
- The mean is a computed average, whereas the median and the mode are positional.
- Extreme values in the series affect the utility of the mean, but not of the median or the mode.
- The presence of open-ended classes excludes the use of the mean, but not of the median and sometimes of the mode.
- Varying class intervals usually make the mean unreliable, but do not affect the median. In such a case, the mode may be found but only by reclassification through frequency densities.
- Means may be combined, but not medians and modes.
- The mean has four mathematical properties which make it indispensable in advanced statistical work. The median has one mathematical property, while the mode has none.
- To compute the median and the mode, the data has to be sorted.
- The median and the mode may be found graphically, but not the mean.

Check Your Progress - 1

1. The median is
   a. 1st Quartile
   b. 2nd Quartile
   c. 50th Percentile
   d. Both (a) and (c) above
   e. Both (b) and (c) above.
2. Which of the following is not a requisite of a good average?
   a. Easy to compute
   b. Capable of simple interpretation
   c. Dependent on all the observed value
   d. Fluctuate relatively more
   e. Capable of mathematical manipulation
3. Identify the option which is not a type of common average
   a. Arithmetic Mean
   b. Geometric Mean
   c. Standard Deviation
   d. Median
   e. Mode
4. Fill-in in the blank in the following statement.
   …….is the most appropriate average in dealing with qualitative data

2.9 Geometric Mean

Managers, sometimes come across quantities that change over a period of time. They might need to know the average rate of change over a period of time. Arithmetic mean proves to be in accurate in this situation. For example, suppose the sales of a company were Rs.10 crore in March, 20x1. By March, 20x2 it increased by 100% to Rs.20 crore. Further, in the next one year it decreased by 50% so that in March, 20x3 the sale was again Rs.10 crore. Hence, the growth rate over the two years is zero. If we calculate the arithmetic mean then the result will be 25% i.e., [(100 + (–50))/2]. This is considered meaningless. So, we can calculate the geometric mean in this situation. Geometric Mean is defined as the nth root of the product of numbers to be averaged. The geometric mean of numbers X_1, X_2, X_3.....X_n is given as: $G = (X_1 \text{ x } X_2 \text{ x } X_3 \ldots\ldots X_n)^{1/n}$.

For the above example, we have to calculate the quantity ratio first:

$$\text{Quantity ratio} = \frac{\text{Percentage growth rate}}{100} + 1$$

So,

$$q_1 = \frac{100}{100} + 1 = 2 \quad \text{and } q_2 = \frac{-50}{100} + 1 = 0.5$$

So,

Geometric mean = $(2 \text{ x } 0.5)^{1/2} = 1^{1/2} = 1$

Average growth rate = Geometric mean –1 = 1 – 1 = 0.

Growth over time is peculiarly due to compounding. For example, sales of a company were Rs.10 crore in 2017-18. In 2019-2020, they grew by 10% to be Rs.11 crore. In 2020-2021, they grew by 20% to be Rs.13.2 crore. The 20% growth rate applies to Rs.11 crore which includes the Rs.1 crore growth of the previous year. This is called ‘compounding’.

Property

The product of the quantity ratios will remain unchanged when the value of geometric mean is substituted for each individual value. This may be seen by substituting 0% for 100% and –50% in the above example.

Uses

The geometric mean is used to find the average percent increase in sales, production, population or other economic or business series overtime.

Example 8

The following data relates to ABC Ltd.

Year	EPS Rs.
2017-2018	6
2018-2019	7.5
2019-2020	12.5

The growth rate for the year 2018-19 = $\frac{1.5}{6}$ x 100 = 25%

The growth rate for the year 2019-20 = $\frac{5}{7.5}$ x 100 = 66.67%

We can find that the EPS of ABC Ltd., has been increasing year by year, but at different growth rates. Now, the compounded annual growth rate can be arrived at by taking the geometric mean for the two quantity ratios.

G.M. $= \sqrt{1.25 \text{ x } 1.67} = 1.44$

Growth rate = 1.445 – 1 = 0.445 or 44.5%

Thus, the compounded annual EPS growth rate of ABC Ltd., for the years 2017-18, 2018-19 and 2019-20 is 44.5%. Geometric mean is most frequently used in finding out compound interest. The formula for compound interest is:

$$P_n = P_0(1 + r)^n$$

Where,

P_n = The value at the end of period n,

P_0 = The value at the beginning of the period,

r = Rate of compound interest per annum (expressed as a fraction),

n = Number of years.

Merits and Limitations of Geometric Mean

Merits

- It is rigidly defined.
- It is useful in average ratios and percentages and in determining rates of increase and decrease.
- It gives less weight to large items and more to small ones than the arithmetic average as it is never larger than the arithmetic mean.
- It is capable of algebraic manipulation.

Limitations

- It is difficult to compute and interpret and so has restricted application.
- It cannot be computed when there are both negative and positive values in a series or one or more of the values are zero.

2.10 Harmonic Mean

The Harmonic Mean is based on the reciprocals of numbers averaged. It is defined as the reciprocal of the arithmetic mean of the given individual observations. If there are N samples each of size n, then the harmonic mean is defined as:

$$\text{HM} = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \ldots\ldots + \frac{1}{X_n}}$$

Where,

X_1, X_2, X_3, etc., refer to various items of the variable.

Weighted Harmonic Mean

Weighted Harmonic Mean is calculated with the help of the following formula:

$$\text{WHM} = \frac{\sum W}{\sum (W/X)}.$$

2.10.1 Appropriateness of the Harmonic Mean, Geometric Mean and Arithmetic Mean

i. For a set of ratios which has been calculated with the same denomination, it is appropriate to use the arithmetic mean.

Example 9

Consider a company consisting of only two divisions X and Y. The calculation of the gross profit ratio for the two divisions as well as for the company as a whole is shown below:

	Division X	Division Y	Whole Company (X + Y)
Gross Profits (Rs. crore)	60	75	135
Sales (Rs. crore)	400	400	800
Net Profit Margin (%)	15	18.75	16.875

Here, we see that the net profit margin for the company as a whole is 8.75% which is nothing but the Simple Arithmetic Mean of the net profit margins of the two divisions A and B:

$$\frac{(15+17.75)}{2} = 16.375$$

So the simple arithmetic mean has significance here because the denominator for both the ratios is 400.

ii. For a set of ratios which has been calculated with the same numerator it is appropriate to use the harmonic mean.

Example 10

Consider a company consisting of only two divisions, X and Y. The calculation of the net profit margins for the two divisions as well as for the company as a whole is given below:

	Division X	Division Y	Whole Company (X + Y)
Gross profits (Rs. crore)	75	45	120
Sales (Rs. crore)	300	450	750
Net Profit Margin	25	10	16.6

Here, we find that the simple arithmetic mean of the net profit margins of the two divisions is (25 + 10)/2 = 17.5%. This is not equal to the net profit margin for the whole company which is 16.6%. Suppose we calculate the harmonic mean of two divisions.

Divisional Gross Profit Margins = $2/\left(\frac{1}{25}+\frac{1}{10}\right)$ = 100/7 = 14.3 (approx)

So, here, the gross profit margin for the whole company is the simple harmonic mean of the gross profit margins of the company's two divisions.

iii. For a set of ratios with different denominators and numerators it is appropriate to use the weighted mean.

Example 11

Consider a company consisting of only two divisions X and Y. The calculation of the gross profit margins for the two divisions as well as for the company as a whole is given below.

	Division X	Division Y	Whole Company (X + Y)
Net profits (Rs. crore)	40	60	100
Sales (Rs. crore)	200	600	800
Net Profit Margin (%)	20	10	12.5

Here, we find that the simple arithmetic mean of the net profit margins of the two divisions is (20 + 10) 2 = 15. The simple harmonic mean of the net profit margins of the two divisions is: $= \dfrac{2}{\dfrac{1}{20}+\dfrac{1}{10}} = 13.33\%$

In this case, both simple harmonic mean and arithmetic mean are not appropriate. So we use the weighted harmonic mean.

Here, if gross profit is used as a weight, the Weighted Harmonic Mean:

$$\text{WHM} = \frac{4+6}{\dfrac{4}{20}+\dfrac{6}{10}} = 12.5\%$$

Which is the gross profit margin for the company as a whole. On the other hand, if we use the denominators (sales) as weights, the appropriate mean is the Weighted Arithmetic Mean (WAM). So,

$$\text{WAM} = \frac{200 \text{ x } 20 + 600 \text{ x } 10}{200+600} = 12.5\%$$

Which is the gross profit margin for the company as a whole.

Note: The above analysis is not valid for ratios covering growth rates over long periods of time where the appropriate mean is the Geometric Mean.

2.11 Measures of Dispersion

In addition to measures of central tendency, it is often desirable to consider measures of variability, or dispersion. For example, suppose you are an investor in the stock market and want to purchase a banking stock. You have selected two fundamental strong banks A and B and calculated the historical mean of this return. Historical mean of these two companies is equal. In this case, which company will you select? Do the two companies have the same degree of reliability in terms of future return? Note the dispersion, or variability, in the return. Which company would you prefer? Your answer will be the company which has low variability in return. The variability or dispersion of data is given by the measures of dispersion. When there is no dispersion, all the data points have identical values and the values of all the measures of central tendency converge.

2.11.1 Range

Range is the simplest measure of the dispersion and is defined as the difference between the highest value data point and the lowest value data point. It can be given as: Range = Highest value data point – Lowest value data point.

Merits and Limitations

Merits

i. Range is simple to understand and easy to compute.

ii. It gives a quick picture of variability, as the time required to calculate is minimum.

Limitations

i. Range does not give the accurate picture of variability as it does not consider each and every item of the distribution.

ii. It fluctuates from sample to sample.

iii. It cannot disclose the character of the distribution within the two extremes.

Despite these limitations, the range is widely used in the following areas:

a. For quality control.

b. In studying the fluctuations in the prices of stocks and shares.

c. In forecasting the weather.

Deviation Measures

Range considers only two data points. One way to overcome this problem is to consider the deviation from every data point, in other words, calculate the difference of all data points from one single point.

2.11.2 Mean Absolute Deviation

Sometimes, to avoid the problem of positive and negative deviations canceling out each other, we can use the Mean Absolute Deviation given by:

$$\frac{\sum\left|X-\bar{X}\right|}{n}.$$

Here,

$$|X-\bar{X}| = X-\bar{X} \text{ if } X \geq \bar{X} \qquad |X-\bar{X}| = \bar{X}-X \text{ if } X \leq \bar{X}$$

However, even though mean deviation avoids the problem of positive and negative deviations canceling out, it is not popular in practice, because it is very difficult to deal with the value $|X - \bar{X}|$. In addition, it does not show the direction of deviation i.e., positive or negative.

Example 12

Income (Rs.)	Deviations from Mean	Absolute Deviation
3,000	3,000 – 3,400= – 400	$\|-400\|$ = 400
3,200	3,200 – 3,400 = – 200	$\|-200\|$ = 200
3,400	3,400 – 3,400 = 0	$\|-0\|$ = 0
3,600	3,600 – 3,400 = 200	$\|200\|$ = 200
3,800	3,800 – 3,400 = 400	$\|400\|$ = 400
Total		1,200

Average mean absolute deviation = 1,200/5 = 240

Merits and Limitations

Merits

i. It is simple to understand and easy to compute.

ii. It is a better measure of dispersion when compared to range and quartile deviation because it includes all observations in its calculation.

iii. Mean deviation is less affected by the extreme values when compared to standard deviation.

iv. Mean deviation is considered as a true and accurate measure of dispersion.

Limitations

i. It considers the absolute value of the deviations and ignores the algebraic signs, which is mathematically unsound and illogical.

ii. Further mathematical treatment is not possible due to the first limitation.

iii. It is rarely used in sociological sciences.

iv. It does not give an accurate result because the deviations are taken from median which will not give satisfactory results in case of high degree of variability.

Despite the above limitations, the mean deviation has wider practical utility in economics and business statistics. It is popular because of its simplicity and accuracy of computations.

2.11.3 Other Measures of Dispersion

Some other measures are interfractile range, interquartile range and quartile deviation. Quartile deviation is also referred to as semi-interquartile range.

Interfractile Range: Interfractile range is the difference between the values of two fractiles. Fractiles are similar to percentages. Depending on the number of equal parts into which we divide the data we call them as deciles, percentiles. A decile is a fractile which divides the data into ten equal parts, a percentile is a fractile which divides the data into 100 equal parts and finally a quartile divides the data into four equal parts. While computing fractiles, we ought to arrange the elements in an increasing order.

Interquartile Range: Interquartile range is the difference between the last value in the third quartile (usually denoted by Q_3) and the last value in the first quartile denoted by Q_1.

Quartile Deviation: It is the difference between the third and the first quartile divided by two. It is expressed as QD:

$$QD = (Q_3 - Q_1)/2$$

Where,

Q_1 = first quartile,

Q_3 = third quartile.

Example 13

The following table gives the distribution of monthly incomes of 500 workers in a factory. Calculate quartile deviation and its coefficient.

Monthly Income (Rs.)	No. of Workers
Below Rs.100	10
100-150	25
150-200	145
200-250	220
250-300	70
300 and above	30

Solution

Calculation of Quartile Deviation and its Coefficient

Monthly Income (Rs.)	No. of Workers (f)	cf
Below – 100	10	10
100-150	25	35
150-200	145	180
200-250	220	400
250-300	70	470
300 and above	30	500

$$Q_1 \;= \text{Size of} \left(\frac{N}{4}\right)^{th} \text{item} = \frac{500}{4} = 12_{th}\text{ h item}$$

$$Q_1 \;= L + \frac{N/4 - cf}{f} \times i \;=\; 150 + \frac{125-35}{145} \times 50 \;= 150 + 31.03 = 181.03$$

$$Q_3 \;= \text{Size of} \left(\frac{3N}{4}\right)^{th} \text{item} = \frac{3(500)}{4} = 3^{th}\text{ th item}$$

$$Q_3 \;= L + \frac{3N/4 - cf}{f} \times i \;=\; 200 + \frac{375-180}{220} \times 50 \;= 200 + 44.32 = 244.32$$

$$QD \;= \frac{Q_3 - Q_1}{2} \;= \frac{244.32 - 181.03}{2} = 31.645$$

2.11.4 Standard Deviation

The main drawback of the different measures of dispersion, as discussed earlier, is that the positive and negative deviations cancel out each other. Use of the mean absolute deviation overcomes this problem, but creates another. It is difficult to deal with a quantity like $\left|X - \bar{X}\right|$ algebraically. To overcome these problems another measure called standard deviation is more popular and frequently used in statistics. To convert the negative number into a positive number, take the square of the negative number. Standard deviation is the square root of mean square deviation. This method offers the following advantages: (i) Squaring makes each negative value into a positive value. Thus, the value below the mean cannot cancel the value above the mean. (ii) By squaring the mean deviation, the unit changes into a square. If the data is expressed in rupees, then when we square the deviations, the units are Rs^2. By taking the square root of the squared deviations, we get the standard deviation. (iii) Squaring of the data gives more weightage to the high value, which is appropriate for statistical measures.

Procedure for calculating standard deviation: (i) Calculation of deviations of the observations from the mean. (ii) Squaring each deviation. (iii) Finding the mean of the squared deviations obtained in step (ii). (iv) Taking the positive square root of the mean found in step (iii). Finally, population standard deviation can be expressed using the formula:

$$\sigma = \frac{\sqrt{\sum (X - \mu)^2}}{N}$$

Where,

σ	–	denotes the population standard deviation.
X	–	denotes each observation.
μ	–	is the arithmetic mean of the population.
N	–	is the number of observations.

The sample standard deviation can be calculated using the sample mean notation. Since sample is an unbiased estimator of population mean, the same cannot be true for sample standard deviation. This can be corrected by using N – 1 in place of N. Thus, s = $\frac{\sqrt{\Sigma(X-\mu)^2}}{N-1}$

Where,

s denotes sample standard deviation,

X denotes each observation,

μ is the arithmetic mean of the population,

N is the number of observations.

2.11.5 Grouped Data

For grouped data, the formula applied is:

$$\sigma = \sqrt{\frac{\Sigma_f (x-\mu)^2}{\Sigma_f}}$$

Where,

f = frequency of the variable,

μ = population mean.

Alternative Method

Sometimes, step deviation method is used for calculating standard deviation from a grouped data. The deviations of mid-points from an assumed mean are taken and divided by the width of the class interval, i.e. 'i'. The formula applied is:

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times i$$

Where,

d = (m – A)/i

i = Class interval,

A = Assumed mean,

m = Mid-point of the class interval.

2.11.6 Variance

Instead of standard deviation sometimes we use variance for calculating variability. Variance is the average squared deviation from the arithmetic mean.

Variance of any data is the square of the standard deviation. It is denoted by σ^2 in case of a population and s^2 in case of a sample.

Example 14

Compute the variance and standard deviation of a population from the following details:

Deviation of class mid-point from assumed mean	Number of Observations
–20	5
–10	10
0	20
10	30
20	20
30	15
Total number of observations	100

The width of the class interval is 10. The variance of the population using assumed mean can be calculated as:

$$\sigma^2 = \left[\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2\right] \times i^2$$

Deviations	$d = \frac{\text{Deviation}}{i}$	Frequency	d^2	fd^2	fd
–20	–2	5	4	20	–10
–10	–1	10	1	10	–10
0	0	20	0	0	0
10	1	30	1	30	30
20	2	20	4	80	40
30	3	15	9	135	45
		100		275	95

Variance of the population = $\left[\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2\right] i^2$

$$= \left[\frac{275}{100} - \left(\frac{95}{100}\right)^2\right] \times (10)^2 = \quad 184.75.$$

Standard Deviation = $\sqrt{\text{Variance}}$ = $\sqrt{184.75}$ = 13.592

Properties

i. Standard deviation is independent of change of origin, but dependent on the change of scale. It means, if in a series each observation is increased or decreased by a constant quantity, the standard deviation will remain the same, but if each observation is multiplied or divided by a constant quantity, standard deviation will also be similarly affected.

ii. The sum of the squares of the deviations of items of any series from a value other than the arithmetic mean would always be greater. In statistical terms, standard deviation is the minimum root-mean-square deviation.

iii. Just as it is possible to compute combined mean of two or more groups, it is also possible to compute combined standard deviation of two or more groups. Combined standard deviation denoted by $\sigma_{1,2}$ is computed as follows:

$$\sigma 12 = \sqrt{\frac{N_1 \sigma_1^2 + N_2 \sigma_2^2 + N_1 d_1^2 + N_2 d_2^2}{N_1 + N_2}}$$

Where,

μ_1 = Mean of first group,

μ_2 = Mean of second group,

σ_1 = Standard deviation of first group,

σ_2 = Standard deviation of second group,

N_1 = Number of observations in the first group,

N_2 = Number of observations in the second group.

d_1 = $\mu_1 - \mu$

d_2 = $\mu_2 - \mu$

μ = $(N_1 \mu_1 + N_2 \mu_2)/(N_1 + N_2)$.

Example 15

Consider the following two sets of observations:

For set 1, $\mu_1 = 10$, $\sigma_1 = 2.5$, $N_1 = 5$

For set 2, $\mu_2 = 15$ $\sigma_2 = 2$, $N_2 = 4$

(Consider these two sets of observations as populations).

Calculate the standard deviation for the combined set of observations.

Combined mean μ

$$= \frac{N_1 \mu_1 + N_2 \mu_2}{N_1 + N_2} = \frac{(5 \text{ x } 10) + (4 \text{ x } 15)}{5 + 4} = 110/9 = 12.22$$

$$d_1 = \mu_1 - \mu = 10 - 12.22 = -2.22$$

$$d_2 = \mu_2 - \mu = 15 - 12.22 = 2.78$$

Standard deviation for the combined set of 9 observations:

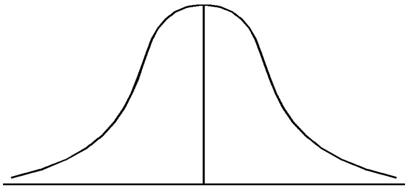
$$= \sqrt{\frac{5\,(2.5)^2 + 4(2)^2 + 5\,(-2.22)^2 + 4(2.78)^2}{5+4}}$$

$$= \sqrt{\frac{5 \text{ x } 6.25 + 4 \text{ x } 4 + 5 \text{ x } 4.92 + 4 \text{ x } 7.7284}{5+4}} = 3.37$$

2.12 Standard Deviation and Normal Curves

Mean and standard deviation allows the statistician to extract useful information. This information depends on the shape of the curve. Under a symmetrical curve, bell shaped curve can be used for the empirical rule.

Figure 2.2: Normal Curve



According to the empirical rule:

- In a normal curve, 68.27% of items in the distribution fall between the range of the arithmetic mean plus and minus 1 standard deviation $(\mu \pm 1\sigma)$.
- 95.45% of the items fall between the arithmetic mean plus and minus 2 standard deviations $(\mu \pm 2\sigma)$.
- 99.73% of the items fall within the range of arithmetic mean plus and minus 3 standard deviations $(\mu \pm 3\sigma)$.

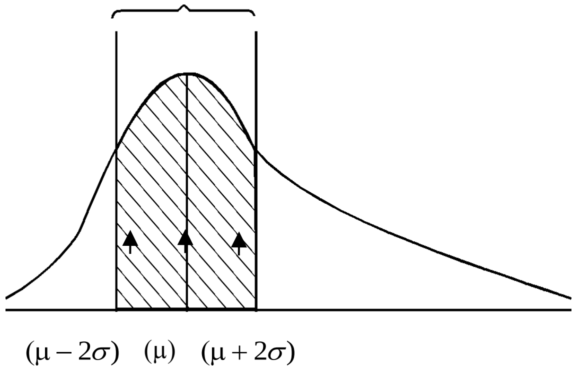
2.13 The Bienayme-Chebyshev Rule

The Bienayme-Chebyshev theorem states that the percentage of data observations lying within $\pm k$ standard deviations of the mean is at least:

$$\left(1 - \frac{1}{k^2}\right) \text{x } 100$$

For example, if k = 2, the value of $\left(1 - \frac{1}{k^2}\right)$ x 100 will be 75%. It means that at least 75% of the observation falls within the ± 2 standard deviation. The figure given below is of a non-symmetrical distribution.

Figure 2.3: Non-symmetrical Distribution



While the Bienayme-Chebyshev rule is applicable to any kind of distribution, the empirical rule is applicable to only symmetrical distribution. This rule helps managers to understand the information content of data when no known distribution can be assumed.

2.14 Application of Standard Deviation in Finance

Standard deviation is normally used to calculate the risk i.e., variability of return in any particular securities or portfolio. Probability distribution of return is used to calculate the standard deviation.

Example 16

The EPS of a group of companies have the following frequency distribution for 20x1-20x2.

EPS (Rs. crore)	Frequency
0-2	5
2-4	8
4-6	7
Total	20

This distribution can be converted into probability distribution using relative frequency approach.

EPS Rs. crore	Mid value	Probability
0-2	1	5/20 = 25%
2-4	3	8/20 = 40%
4-6	5	7/20 = 35%
Total		100%

Note that the total probabilities (which represent the likelihood of all possible outcomes) always equal 100% or 1.

Expected EPS will be = 1 x 0.25 + 3 x 0.4 + 5 x 0.35 = 3.2

The standard deviation of a variable is calculated using the following formula:

$$\sigma = \left[\sum_{i=1}^{n} p_i (k_i - \bar{k})^2\right]^{1/2}$$

Where,

p_i	=	Probability associated with the occurrence of ith rate of return,
k_i	=	ith possible rate of return,
$\bar{k}$	=	Expected rate of return, i.e., mean,
n	=	Number of possible outcomes.

Let us calculate the standard deviation of EPS.

EPS	$E_i - \bar{E}$	$\left(E_i - \bar{E}\right)^2$	p_i	$\left(E_i - \bar{E}\right)^2 x\, p_i$
1	–2.2	4.84	0.25	1.21
3	–0.2	0.04	0.40	0.016
5	0.8	0.64	0.35	0.224
				1.41

$\bar{E}$ is expected EPS and is the probability weighted average of possible outcomes of the random variable here. Expected Return:

$$\bar{E} = \Sigma\, p_i\ E_i = 1\,x\,0.25 + 3\,x\,0.40 + 5\,x\,0.35 = 3.2$$

$$\sigma = \left[\sum_{i=1}^{n} p_i (k_i - \bar{k})^2\right]^{1/2} = \sqrt{1.41} = 1.187$$

Risk of Portfolios

Standard deviation is frequently used in the calculation of portfolio risk, which is calculated to know-how much risk has been reduced through it. A brief discussion of the calculation of portfolio risk follows in the chapter, “Linear Regression”.

2.15 Coefficient of Variation

Sometimes, selection of securities for investment based on historical mean return creates a problem. Suppose there are two securities A and B. Their average return and standard deviation are given below.

Stock	A	B
Mean Return	20%	20%
Standard Deviation	3.26%	5.25%

In the above example, return on both the stock is same. An investor will choose the stock which has the lowest risk i.e., stock A. Lower standard deviation represents the lower variation in price movement of stock and hence less risk is involved. Though standard deviation discussed above is an absolute measure, coefficient of variation is a relative measure and based on the standard deviation.

$$\text{Coefficient of variation} = \frac{\text{Standard Deviation}}{\text{Mean}} \times 100$$

It is more appropriate statistical measure than standard deviation, to compare the variability, homogeneity, stability, uniformity and consistency of two or more series. Higher coefficient of variation represents more variability or conversely less consistency, less uniformity and lower coefficient of variation represents less variability or more consistency, more uniformity, more stability or more homogeneity.

Example 17

For example, coefficient of variation for the above discussed example is as follows:

Stock	A	B
Mean Return	20%	20%
Standard Deviation	3.26%	5.25%
Coefficient of Variation	16.3%	26.25

Since, stock B has higher Coefficient of Variation, so the stock B is more risky. In other words, future cash flow of B is less consistent compared to stock A.

Example 18

Q. Students who take admission in a business school are from diverse academic background (e.g. art, commerce, science, and engineering). Some of the courses that are taught at these schools require knowledge of mathematics. To assist the students of non-science and non-engineering background, some of these business schools run preparatory courses. Teachers from a business school want to assess the performance of a group of students who have attended special mathematics preparatory classes for students from a non-mathematics background. The data from a sample of 10 students is given below. Calculate the values of different measures of central tendency (mean, median, mode) for this data.

S No	Marks
1	5
2	10
3	13
4	17
5	8
6	12
7	6
8	16
9	15
10	7

Solution:

Mean marks: 10.9

Median marks: 11

Mode marks: None (as there is no value which occurs more than once).

Example 19

Q. Goa is one of the popular tourist destinations in India. Many tourists, who visit Goa, usually take self-driven private vehicles for rent. Cerejo, a resident of Panjim, is running the vehicle renting business, to supplement his income. He predominantly follows two business models that are quite prevalent in this business. In first, there is a nominal fixed charge for renting of vehicle for a given duration. Here, the cost of petrol is to be borne by the user. In the second model, a lumpsum amount is charged for a given duration, also covering the cost of petrol. To find the fluctuation in utilization of a particular model of SUV that he rents, he wants to assess the variability in petrol used in that vehicle for the last 12 months.

a) Use the following data to help him calculate the range, variance, and standard deviation for monthly consumption of petrol.

b) If Cerejo wants to compare the consumption variability of the SUV with a motorcycle that he rents, what measure(s) should he use?

Month	Consumption
1	100
2	120
3	130
4	150
5	250
6	240
7	220
8	160
9	180
10	140
11	300
12	200

Solution:

a) Range: 200 litres

Variance: 3656.82 litres

Standard Deviation: 60.47

b) The above measures (range, variance, and standard deviation) may not be appropriate for comparing the variability in two different contexts (here an SUV with a motorcycle). Coefficient of variation that takes into consideration the context would be a more appropriate measure. Its value here (in terms of percentage) is 33.14%.

Check Your Progress - 2

5. Which of the following is not a measure of dispersion?

   a. Mean

   b. Range

   c. Mean Absolute Deviation

   d. Standard Deviation

   e. Quartile Deviation

6. Fill in the blank with appropriate words in the following statement

   Averaged squared deviation from the arithmetic mean is called ……..

7. The growth rates of a textile unit in the western region for the last five years are given below:

Year :	1	2	3	4	5
Growth Rate :	7	8	10	12	18

   What is the geometric mean of the growth rate?

   a. 1.266.

   b. 1.1093.

   c. 1.1.

   d. 1.31.

   e. 1.42.

8. A data set includes some quantities. The sum of reciprocals of the quantities is 7/8. The harmonic mean of the data set is 24/7. The data set is expanded by including the quantities 5 and 10.

   What is the harmonic mean of the expanded data set?

   a. 7/24

   b. 10 2/21

   c. 6 2/3

   d. 7 3/7

   e. 4 12/47

9. For the given set of collected data, what is the range?

   12,32,21,23,34,32,23,34,31,21,23,4,43,42,24,3,2,32,43,32,23,42,34,52

   a. 48
   b. 24
   c. 40
   d. 44
   e. 10

10. The following details are available with regard to two groups of data, A & B:

Group	Number of Observations	Mean	Standard Deviation
A	15	20	4
B	25	16	2

   What is the combined standard deviation for both the groups?

   a. 2.50
   b. 3.53
   c. 6.25
   d. 7.00
   e. 12.25

2.16 Summary

- Central tendency and the dispersion are two vital parameters frequently used in statistics. The measure of central tendency is a single value that is used to represent whole data set. Central tendency, can be measured mainly through three ways, i.e., mean, median and mode. Mean is the average of all the observations, median measures the mid value in the observation and mode is the value most often repeated in the data set. All these measures are have their own advantages and disadvantages and are used at different situations.
- Dispersion measure how much data in a given set of numerical data is spread out. Range is the simplest measure of dispersion and is defined as the distance between the highest and the lowest values in a data set. The other measures are quartile deviation, mean deviation and standard deviation. The standard deviation is the best measure of dispersion and widely used in various areas to measure risk.

2.17 Glossary

Dispersion: A degree of variation about a central value.

Geometric Mean: A measure of central tendency used to measure the average rate of growth, computed by taking the ‘n’th root of the product of n values representing change.

Harmonic Mean: It is the reciprocal of the arithmetic mean of the reciprocal of the given individual observations.

Mean Deviation: The sum of deviation from an average divided by the number of units.

Median: The middle point of a data set, or a measure of location that divides the data set into halves.

Mode: The value most often repeated in the data set.

Range: The distance between the highest and lowest values in a data set.

Standard Deviation: The square root of mean square deviation.

Variance: The square of standard deviation.

2.18 Suggested Readings/Reference Material

1. Gupta, S. P. Statistical Methods. 46th Revised ed. New Delhi: Sultan Chand & Sons. 2021.
2. I. Levin Richard, H. Siddiqui Masood, S. Rubin David, Rastogi Sanjay. Statistics for Management. Pearson Education; Eighth edition, 2017.
3. Gerald Keller. Statistics for Management and Economics. Cengage, 2017.
4. Arora, P. N., and Arora, S. CA Foundation Course Statistics. 6th ed. S Chand Publishing, 2018.
5. Mario F Triola. Elementary Statistics. 13th ed., 2018.
6. David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran. Statistics for Business and Economics. 13th Edition, Cengage Learning India Pvt. Ltd., 2019.
7. S D Sharma. Operations Research. Kedar Nath Ram Nath, 2018.
8. Hamdy A. Taha. Operations Research: An Introduction. 10th ed., Pearson, 2016.

9 Malhotra, N. (2012), Marketing Research: An Applied Orientation, 7th ed., Pearson, 2019.

10. Cooper, D.R. and Schindler, P.S. and J. K. Sharma (2018), Business Research Methods, 12th edition, McGraw-Hill Education.

2.19 Self-Assessment Questions

1. Describe the appropriateness of three measure of central tendency.

2. What are the requisites of a good average? Explain.

3. What do you understand by dispersion? What are the different measures of dispersion?

4. What is meant by coefficient of variation? Explain the difference between variation and mean deviation.

5. Explain the term “quartile deviation”.

2.20 Answers to Check Your Progress Questions

1. (c). 50th percentile

2. (d). Fluctuate relatively more

3. (c) Standard Deviation

4. Median

5. Mean

6. Variance

7. (a) 1.266

8. (e) 4 12/47

9. (a) 48

10. (b) 3.53

Unit 3

Probability

Structure

3.1 Introduction

In the previous unit, you studied concepts of different types of measures of central tendencies – mean, median, and mode and measures of dispersion – range, variance, standard deviation, etc. In this unit, you will learn the concept of probability and its application in statistics. Management decisions are made for the future of a firm. But the future is uncertain, things may happen or may not happen; this uncertainty can be reduced with probability. For example, when you say that there is 90 percent chance for the company’s sales to reach Rs.100 crore next year, it means the company’s sales may reach Rs.100 crore the next year or it will be at least Rs.90 crore (100 X 0.90). Probability in simple terms means the likelihood that an event will occur. It is an important part of statistics.

3.2 Objectives

After going through the unit, you should be able to:

- Explain basic concepts of probability;
- Identify simple and compound events;
- Recognize mutually exclusive events;
- Analyze three conceptual approaches of probability;
- Demonstrate properties of probability;
- Define dependent and independent events;
- Explain unconditional and conditional probability; and
- Use of Bayes’ theorem.

3.3 Some Basic Concepts of Probability

Business managers today have to take some critical decisions based on future expectations. They often want to know the chances of either an increase or decrease in the sales of the company, and the projects that will be finished on time, etc. Probability is defined as the numerical measure of the likelihood that an event will occur. Thus, it could be used as measure of the degree of uncertainty associated with events.

Probability values are always assigned on a scale from zero to one. A probability near zero means an event is unlikely to occur and a probability near 1 indicates almost 100 percent certainly that an event will occur. Let us discuss some basic concepts of probability.

Experiment, Outcomes and Sample Space: Experiment is defined as a process that allows the statistician to make an observation. These observations are called outcomes of the experiments. The sample space S of a random experiment is the set of all possible outcomes of the experiment. The table given below is an example of an experiment, its outcomes and sample space:

Table 3.1: Example of an Experiment, Outcomes and Sample Space

Experiment	Outcomes	Sample Space
Role a dice once	1, 2, 3, 4, 5, 6	S = {1, 2, 3, 4, 5, 6}
Toss a coin once	H T	S = {H T}
Select people	M F	S = {M F}
Purchasing shares of a common stock	Increase, decrease or no change in price per share	S = {Increase, decrease, no change}
Play a football game	Win, lose, or tie	S = {Win, lose or tie}

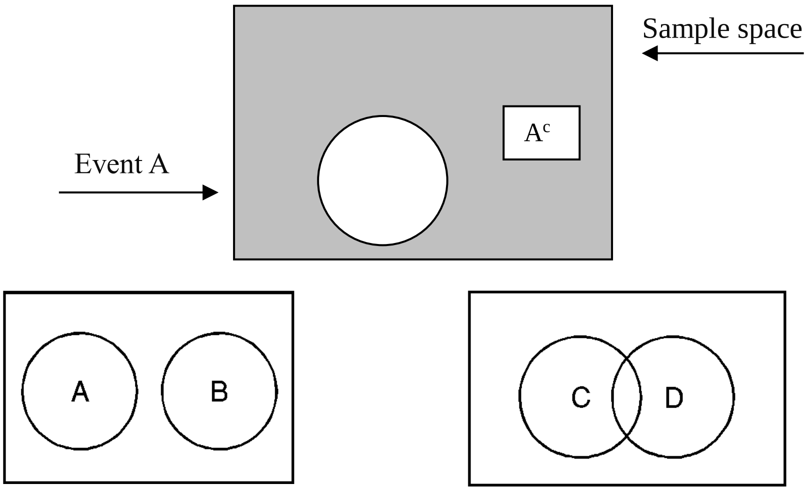
Mutually Exclusive Event: Events A and B are mutually exclusive if, when one event occurs, the other cannot. For example, consider a firm planning to invest in of similar projects say, A, B and C. If the firm decides to invest in project A, it means it has rejected projects B and C. The acceptance of one precludes the acceptance of another. Thus, acceptance of projects A, B and C is said to be mutually exclusive event. Tossing a coin is a mutually exclusive, because head and tail cannot occur at the same time. Similarly, rolling a die is also a mutually exclusive event as the entire outcomes 1, 2, 3, 4, 5, 6, cannot occur at the same time.

Exhaustive Event: Exhaustive events include all possible outcomes of an experiment. In tossing an unbiased coin, the occurrence of head or tail are exhaustive events.

Complement of an Event: For a given event A, the complement of A is defined to be the event consisting of all sample points that are not in A. Complement of A is represented by A^c or $\overline{A}$.

The Venn Diagram: A Venn diagram is a pictorial representation of the sample space of an experiment. Figure 1 is a Venn diagram that illustrates the concept of a complement. The rectangular area represents the sample space for the experiment and as such contains all possible sample points. The circle represents event A and contains only the sample point that belongs to A. The shaded region of the rectangle contains all sample points not in event A, and is by definition the complement of A.

Figure 3.1: Venn Diagram of Mutually and Non-mutually Exclusive Events



A & B are mutually exclusive events

C & D are not mutually exclusive events

Remember that in mutually exclusive events only one of them can take place at a time.

3.4 Three Approaches of Probability

The three basic approaches of probability are:

i. Classical Approach;

ii. Relative Frequency Approach;

iii. Subjective Approach.

3.4.1 Classical Approach

All mutually exclusive events like tossing a coin or rolling a dice are activities associated with a classical approach to probability. Probability of an event under this approach is defined as:

$$\text{Probability of an event} = \frac{\text{Number of outcomes for an event}}{\text{Total number of possible outcomes}}$$

This probability is based on prior information about outcomes. i.e., outcomes known without actual experiment, so, it is called ‘priori probability’.

Example 1

Suppose we roll a fair die. There are six equally likely outcomes: 1, 2, 3, 4, 5 and 6. What is the probability of getting a 6?

$$\text{Probability of getting a 6} = \frac{\text{Number of outcomes for an event}}{\text{Total number of possible outcomes}} = \frac{1}{6}$$

3.4.2 Relative Frequency Approach

Relative frequency approach is useful in finding out:

i. The probability that a randomly selected family owns a house.

ii. The probability that the tossing of an unbiased coin will result in a head.

In the relative frequency approach, probability of an event is the proportion of the number of times an event is observed in a very large number of trials.

$$\text{Probability} = \frac{\text{Number of trials in which the event occurs}}{\text{Total number of trials}} \text{ or } \frac{n}{f}$$

Example 2

20 out of 100 randomly selected cars manufactured in a company are found to be of 250cc. Assuming that the 250cc cars are manufactured randomly, what is the probability that the next car manufactured in that company would be of 250cc?

Here, n = 20; f = 100;

So, $P = \frac{20}{100} = 0.2$

3.4.3 Subjective Approach

Some experiments neither have equal outcomes nor repeatedly generated data. Here, we cannot estimate the probability by using the classical approach or the relative frequency approach. Probability can be calculated using the personal judgment approach instead. For example, what is the probability of a major earthquake at a given place? In this case, there is no evidence relative frequency of occurrence. Experts must judge on the information available. The probability statement depends upon the individual’s experience and his familiarity with the facts of the case.

3.5 Probability Rules

i. The probability of any event is equal to the sum of the probabilities of the sample points in the event. For example, in the tossing of a coin the probability of either a head or tail will be 1. P(S) = 1.

   So, for a complimentary event: $P(A^C) = 1 - P(A)$.

ii. The probability of an event always lies in the range of 0 to 1.

   i.e. $0 \leq P(A) \leq 1$.

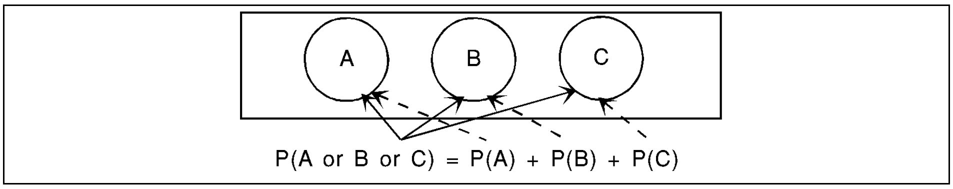
In other words, probability cannot be negative and as the probability of the sample space is 1, the probability of an event contained in the sample space should be less than or equal to 1.

3.5.1 The Addition Rule: Mutually Exclusive Events

When events are mutually exclusive: P(A or B or C) = P(A) + P(B) + P(C).

This can be represented by the Venn diagram as follows:

Figure 3.2: A, B and C are Mutually Exclusive Events



Example 3

A die if thrown once, what is the probability of getting 4 or more?

Let,	A	=	Getting 4 on a throw of a die,
	B	=	Getting 5 on a throw of a die,
	C	=	Getting 6 on a throw of a die.

As there are six possible equally likely outcomes on throwing a die,

$$P\ (A \text{ or } B \text{ or } C) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = P\ (A) + P\ (B) + P\ (C).$$

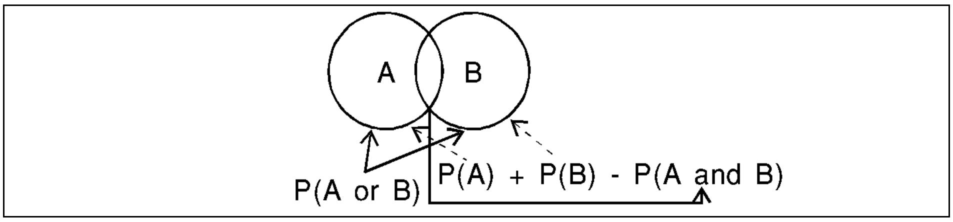
3.5.2 The Addition Rule: Non-mutually Exclusive Events

In this case it is possible for two probabilities to occur at a time. So, while calculating the probability of one of them occurring, the common part i.e., joint probability must be deducted. In other words:

P(A or B) = P(A) + P(B) – P(A and B)

A Venn diagram illustrating the above is given below:

Figure 3.3: A and B are Non-mutually Exclusive Events



Note: P(A or B) is also called union of two events and can be written as P(A ∪ B) and P(A and B) is also called intersection of two events and can be written as P(A ∩ B). P(A and B) or P(AB) is also known as joint probability of two events.

Remember:

$$\boxed{P(\overline{A} \text{ and } \overline{B}) = 1 - P(A \text{ or } B)}$$

Example 4

If P(A and B) = $\frac{1}{2}$ $P(\overline{A} \text{ and } \overline{B}) = \frac{1}{3}$ and P(A) = P(B) = P, then, what is the value of P, if $\overline{A}$ and $\overline{B}$ are complements of events A and B respectively?

We have,

$P(A \text{ and } B) = 1/2$ and $P(\overline{A} \text{ and } \overline{B}) = 1/3 = 1 - P(A \text{ or } B)$

or, $P(A) + P(B) - P(A \text{ and } B) = 1 - \frac{1}{3} = 2/3$

$\therefore$ $P(A) + P(B) = P(A \text{ or } B) + P(A \text{ and } B) = 2/3 + 1/2 = 7/6.$

Now, P(A) = P(B) = P. Therefore, $2P = \frac{7}{6}$ or P = 7/12.

3.5.3 Marginal or Unconditional Probability

Marginal or unconditional probability is defined as the probability where only one event can take place. Marginal probability of an event A is denoted by P(A) and can be calculated using the ratio of a possible event favoring A by the total number of possible outcomes.

$$P(A) = \frac{\text{Number of possible outcomes favoring A}}{\text{Total number of possible outcomes}}$$

The definition assumes that the elements of the sample space have an equally likely chance of occurring.

Example 5

Supposes a company has 1000 employees out of which 500 males are below 40 years of age and 200 are above 40 years; 100 females are below 40 years of age and 200 are above 40 years. The above information can be summarized as follows:

	Below 40 years	Above 40 years	Total
Male	500	200	700
Female	100	200	300
Total	600	400	1000

In the above table, employees can be classified on the basis of gender and age. The probability of each of these four characteristics or events is called the marginal probability. For example, one employee is selected at random from 1000 employees. What is the probability of a male being selected?

$$P(M) = \frac{\text{Number of possible outcomes favoring male}}{\text{Total number of employees}} = \frac{700}{1000} = 0.7$$

Similarly,

$$P(F) = \frac{300}{1000} = 0.3$$

$$P(\text{above 40 years}) = \frac{400}{1000} = 0.4 \text{ and } P(\text{below 40years}) = \frac{600}{1000} = 0.6$$

3.5.4 Conditional Probability

If we want to calculate the probability of males below 40 years of age, this is called as conditional probability. Conditional probability is the probability that an event will occur, given that another event has already occurred. For two events A and B, if B has already happened, then probability of A is written as P(A/B). In the above discussed example, probability of males below 40 years can be written as: P(M/ below 40).

Example 6

In the example discussed above, calculate the probability of males who are below 40 years.

	Below 40 years	Above 40 years	Total
Male	500	200	700
Female	100	200	300
Total	600	400	1000

Required probability is calculated as follows:

	Below 40 years (Y_1)
Male	500
Female	100
Total	600

$$P(M/\text{above } 40) = \frac{\text{Number of males below 40}}{\text{Total number of employees below 40}} = \frac{500}{600} = 0.83$$

Similarly, $P(F/\text{above } 40) = \frac{100}{600} = 0.17$

Now suppose we have to calculate the probability of males above 40 years.

Required probability is as follows:

	Below 40 years	Above 40 years	Total
Male	500	200	700

$$P(\text{above 40/male}) = \frac{200}{700} = 0.285$$

3.5.5 Independent and Dependent Events

Two events are said to be independent if the occurrence of one does not affect the probability of occurrence of the other. Two events A and B, are said to be independent if: P (A/B) = P (A) or P (B/A) = P(B).

Two events are said to be dependent events if the occurrence of one event say A is related to the occurrence of another event, say B. For dependent events:

$$P\ (A/B) \neq P\ (A) \text{ or } P\ (B/A) \neq P\ (B).$$

Example 7

Consider the Example 5 table:

P (above 40 years/male) = $\frac{200}{700} = 0.285$ and P(above 40 years) = $\frac{400}{1000} = 0.4$;

So, the event "male above 40 years" and "male" Dependent events because:

$$P\ (\text{above 40 years/male}) \neq P(\text{above 40years}).$$

But if you consider the data given in the following table:

	Below 40 years	Above 40 years	Total
Male	420	280	700
Female	180	120	300
Total	600	400	1000

P(above 40 years/male) = $\frac{280}{700} = 0.40$ and P(above 40 years) = $\frac{400}{1000} = 0.4$

So, the event "male above 40 years" and "male" Independent because

$$P(\text{above 40 years/male}) = P(\text{above 40 years}).$$

3.5.6 Multiplication Rule: Independent Events

Joint probability of two independent events is equal to the product of their marginal probabilities.

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

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

We may extend the multiplication rule for independent events to three or more events by the following formula: P(A and B and C) = P(A) P(B) P(C).

Example 8

Two fair dice are thrown. What is the probability that one of them gives an even number less than 5, and the other gives an odd number less than 4?

Probability that the first die gives an even number less than 5 and the second die gives an odd number less than 4 = $\frac{2}{6} \text{x} \frac{2}{6} = \frac{1}{9}$.

Probability that the second die gives an even number less than 5 and the first die gives an odd number less than 4 = $\frac{2}{6} \text{x} \frac{2}{6} = \frac{1}{9}$.

Since these events are mutually exclusive, the probability that one of the dice gives an even number less than 5 and the other gives an odd number less than 4 $= \frac{1}{9} + \frac{1}{9} = \frac{2}{9}$.

3.5.7 Multiplication Rule: Dependent Events

The joint probability of two events A and B which are dependent is equal to the probability of A multiplied by the probability of B given that A has occurred.

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

$$\text{or } P(B \text{ and } A) = P(B)\,P(A|B).$$

Example 9

Consider the table from example 5 dependent events:

Now we have to calculate the joint probability of A and B i.e., P(AB).

P(A) = Probability of male, and

P(A/B) = Probability of male above 40 years.

Since P(A) and P(B) are the dependent events:

$$P(AB) = P(A).\ P(A/B) = 0.7 \text{ x } 0.285 = 0.2$$

Check Your Progress - 1

1. There are 5 green 7 red balls. Two balls are selected one by one without replacement. Find the probability that first is green and second is red.
   a. 5/18.
   b. 1/9.
   c. 10/153.
   d. 35/132.
   e. 8/9.
2. A problem is given to three persons P, Q, R whose respective chances of solving it are 2/7, 4/7, 4/9 respectively. What is the probability that the problem is solved?
   a. 122/147
   b. 32/49
   c. 32/441
   d. 122/441
   e. 32/147

3. Three bags contain 3 red, 7 black; 8 red, 2 black, and 4 red & 6 black balls respectively. 1 of the bags is selected at random and a ball is drawn from it. If the ball drawn is red, find the probability that it is drawn from the third bag.
   a. 1/14.
   b. 9/14.
   c. 5/14.
   d. 4/15.
   e. 1/3.
4. Fill in the blank for the following statement.

   Probability could be used as a measure of the ……….associated with events
5. Events A and B are mutually exclusive if,
   a. A and B can never occur
   b. A occurs, and B cannot occur
   c. Both A and B occur
   d. Another event C occurs along with A and B
   e. Always an external event occurs

3.5.8 Conditional Probability: Dependent Event

Conditional probability is the probability that a second event will occur when the first event had already occurred or a first event will occur when the second event has already occurred. For a dependent event, the conditional probability can be expressed as follows:

If A and B are two events then: P(AB) = P(A). P(A|B)

So, $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$ and $P(B/A) = \frac{P(A \text{ and } B)}{P(A)}$; P(A|B) = P(A).

3.5.9 Conditional Probability: Independent Events

In case of A and B being independent events, the probability of event A, give event B has been occurred is as: P (A|B) = P(A). It is so because independent events are those whose probabilities are in no way affected by the occurrence of each other. Hence the probability of event B, give event A has been occurred is as: P (B|A) = P(B).

3.5.10 Marginal Probability: Independent and Dependent Events

We have already discussed that marginal probability is the simple and unconditional probability of an event. For example, in rolling a die the probability of getting one, two or three is always 1/6. These probabilities will be the same if we roll a die a second or a third time or conduct n number of experiments. Thus, each outcome is independent of each experiment. This is not true for a dependent event, where the marginal probability for dependent events is computed by summing up the probability of all joint events in which the simple event occurs. So, if the event A is dependent on events B and C, the joint probability of A will be:

P(A) = P(AB) + P(AC).

Example 10

From the data summarized below for a dependent event, we can calculate the different types of probabilities:

	Below 40 years (Y_1)	Above 40 years (Y_2)	Total
A	500	200	700
B	100	200	300
Total	600	400	1000

Conditional Probability

$P(A/ Y_1) = 500/700 = 5/7$ $P(Y_1/A) = 500/600 = 5/6$

$P(A/ Y_2) = 200/700 = 2/7$ $P(Y_2/A) = 200/400 = 1/2$

$P(B/ Y_1) = 100/300 = 1/3$ $P(Y_1/B) = 100/600 = 1/6$

$P(B/ Y_2) = 200/300 = 2/3$ $P(Y_2/B) = 200/400 = 1/2$

Joint Probability

$P(A \text{ and } Y_1) = 500/1000 = 0.5$

$P(B \text{ and } Y_1) = 100/1000 = 0.1$

$P(A \text{ and } Y_2) = 200/1000 = 0.2$

$P(B \text{ and } Y_2) = 200/1000 = 0.2$

Here, the denominator is divided by 1000 because the sample space for joint probability is 1000.

Marginal Probability

$P(A) = 700/1000$ or $P(A \text{ and } Y_1) + P(B \text{ and } Y_1)$

$P(B) = 300/1000$ or $P(B \text{ and } Y_1) + P(B \text{ and } Y_2)$

$P(Y_1) = 600/1000 = 0.6$ or $P(A \text{ and } Y_1) + P(B \text{ and } Y_1)$

$P(Y_2) = 400/1000 = 0.4$ or $P(A \text{ and } Y_2) + P(B \text{ and } Y_2)$

The additional problem of $P(A \text{ or } Y_1)$ can be calculated as follows:

$P(A \text{ or } Y_1) = P(A) + P(B) – P(A \text{ and } B) = 0.7 + 0.6 – 0.5 = 0.8$

3.6 Bayes' Theorem

Bayes' theorem states that if the prior information about an event is available, we can use it to learn more about that event. For example, suppose 60% of products are made in factory A, and 40% are made in factory B. For suppose randomly selected products, the probability from factory A is 0.60. Suppose we learn that some products are defective and the defect rate for factory A is 35% and for factory B it is 25%. We can use Bayes' formula to calculate the number of defective products from factory A. Bayes' theorem describes the relationships that exist within an array of simple and conditional probabilities. It gives us a way to calculate P(A|B) from knowledge of P (B|A). The probability of A occurring given that B has already occurred is the posterior (or revised) probability. The events with prior probabilities produce, cause, or precede another event, say B. A conditional probability relation exists between events A_1, A_2, ..., A_k and event B. The conditional probabilities are:

$$P(B|A_1),\ P(B|A_2),\ ...,\ P(B|A_k).$$

For any event A_i, Bayes' theorem is expressed in the form of:

$$P(A_i \mid B) = \frac{P(A_i \text{ and } B)}{P(B)}$$

But we know: $P(A_i \text{ and } B) = P(A_i)\ P(B|A_i)$, Since A_1, A_2,…., A_k forms a part of the entire sample space when event B occurs, only one of the events in the part occurs. Thus, we have: $P(B) = P(A_1 \text{ and } B) + P(A_2 \text{ and } B) + ... + P(A_k \text{ and } B)$.

We already know that for any event A_i,

$$P(A_i \text{ and } B) = P(A_i)\ P(B|A_i)$$

When we substitute the formula for $P(A_i$ and B) in the equation for P(B) we obtain:

$$P(B) = P(A_1)\ P(B/A_1) + P(A_2)\ P(B|A_2) +...+ P(A_k)\ P(B/A_k)$$

If we then substitute P(B) and $P(A_i$ and B) into the conditional probability, i.e.,

$$P(A_i|B) = \frac{P(A_i \text{ and } B)}{P(B)}$$

We obtain the generalized version of Bayes' formula, which is given below:

Bayes' Theorem

$$P(A_i \mid B) = \frac{P(A_i)\ P(B|A_i)}{P(A_1)\ P(B|A_1) + P(A_2)\ P\ (B|A_2) +....... + P(A_k)\ P(B|A_k)}$$

Example 11

An item is manufactured by three machines M_1, M_2 and M_3. Out of the total manufactured number of items during a specified production period, 40% are manufactured on M_1, 20% on M_2 and 40% on M_3. It is also known that 5% of the item, produced by M_1 and M_2 are defective, while 4% of those manufactured by M_3 are defective. All the items are put into one bin. From the bin, one item is drawn at random and is found to be defective. What is the probability that it was made on M_1?

Let,

A : The item is defective

B_1 : The item was made on M_1

B_2 : The item was made on M_2

B_3 : The item was made on M_3.

If an item is drawn at random, we know that the prior probability:

$$P(B_1) = 0.40,\ P(B_2) = 0.40 \text{ and } P(B_3) = 0.40$$

We also know that,

$$P(A|B_1) = 0.05,\ P(A|B_2) = 0.05,\ P(A|B_3) = 0.04.$$

These are the conditional probabilities.

We are interested in finding the posterior probability, i.e., having known that the item is defective, what is the probability that it was manufactured on M_1?

$$P(B_1 \mid A) = \frac{P(A \mid B_1) \text{ x } P(B_1)}{P(A \mid B_1) \text{ x } P(B_1) + P(A \mid B_2) \text{ x } P(B_2) + P(A \mid B_3) \text{ x } P(B_3)}$$

$$= \frac{0.05 \text{x} 0.40}{0.05 \text{x} 0.40 + 0.05 \text{x} 0.20 + 0.04 \text{ x } 0.40}$$

$$= \frac{0.02}{0.02 + 0.01 + 0.016} = \frac{0.02}{0.046} = 0.434$$

Similarly, we can find the probability that it was made on M_2 and M_3.

The following example gives an application of Bayes Theorem in real life situation – on COVID Testing – how prior probability helps in finding out reliable and correct results.

Example: Application of Bayes Theorem on COVID Testing

A person undergoes a COVID test. For a moment, we assume that the test is accurate. For instance, it shows 99% of the time correct results when one has COVID infection. Similarly it gives the correct results 99% of the time, if one doesn’t have the disease.

Suppose the disease is very rare. It means one person among every 10,000 has the disease. This is exactly called ad prior probability.

Contd….

If we test one million people and the test correctly shows that 99 of them have it when actually 100 people have the disease. That means 999,900 people don't have the disease. The test correctly determines 989, 901 of 999,900 people. It means that the test determined that 9999 people are with the disease, whereas those people have not had the disease in the actual sense. In this scenario, if a person is identified that he/she has the disease, he/she has $\frac{99}{9999}$ probability of having the disease in reality.

Imagine the test has been taken at face value, then it would have scared a lot of people and they would have to be treated using potentially dangerous medications because of the misdiagnosis. The moral here is that without knowing the prior probability we may not know how likely it will give correct result. This example shows the prior probability in Bayes theorem. With the prior information or probability about an event that is available, we can use it to learn more about that event or the more reliable results about that event.

Adapted from https://www.theguardian.com/world/2021/apr/18/obscure-maths-bayes-theorem-reliability-covid-lateral-flow-tests-probability (Accessed on: 17.09.2021

Check Your Progress - 2

6. Fill in the blank for the statement of Bay's theorem.

   Bay's theorem states that if the prior information about an event is available, We can use it to ……..

7. Which of the following is correct for Conditional probability?

   a. The probability of one event occurring given that another event has occurred.
   b. Probability estimates made prior to receiving new information
   c. A probability that has been revised after additional information was obtained
   d. The set of all the possible outcomes of an experiment
   e. A list of events that represents the possible outcomes of an experiment

8. Tossing coin or rolling a dice are activities associated with which type of probability?

   a. Conditional probability
   b. Posterior probability
   c. Classical probability
   d. Priory probability
   e. Mutually exclusive events

9. In a class, 40% of the students study math and science. 60% of the students study math. What is the probability of a student studying science given he/she is already studying math?’

    a. 0.15.
    b. 0.30.
    c. 0.595.
    d. 0.912.
    e. 0.67

10. A box of fuses contains 20 fuses, of which five are defective. If fuses are drawn one at a time from the box and are not replaced, what is the probability that three draws will result in three defective fuses?

    a. 1/114.
    b. 113/114.
    c. 15/18.
    d. 15/19.
    e. 1/4.

3.7 Summary

- Probability is simply the numerical measure of the likelihood that an event will occur. The first step in calculating probability is to collect all the outcomes of an experiment. The second step is to decide which approach of probability is suitable for a particular situation. All mutually exclusive events like tossing a coin or rolling a die are activities associated with the classical approach to probability. Relative frequency defines the probability of event A if we repeat the experiment several times under the same or similar conditions.
- Subjective probability is based on the judgment of a person or a group of persons. We can use the addition rule and the multiplication rule to calculate probabilities. Further, probability may be conditional, dependent or independent. In each situation, we can use a suitable formula to calculate probability. Bayes’ Theorem is used to calculate posterior (or revised) probability.

3.8 Glossary

Classical Probability: The number of outcomes favorable to the occurrence of an event divided by the total number of possible outcomes.

Collectively Exhaustive Events: A list of events that represents all the possible outcomes of an experiment.

Conditional Probability: The probability of one event occurring, given that another event has occurred.

Event: An event *A* is defined as a subset of the sample space.

Mutually Exclusive Events: Events that cannot happen together.

Posterior Probability: A probability that has been revised after additional information was obtained.

Priori Probability: Probability estimates made prior to receiving new information.

Sample Space: The set of all the possible outcomes of an experiment.

3.9 Suggested Readings/Reference Material

1. Gupta, S. P. Statistical Methods. 46th Revised ed. New Delhi: Sultan Chand & Sons. 2021.
2. I. Levin Richard, H. Siddiqui Masood, S. Rubin David, Rastogi Sanjay. Statistics for Management. Pearson Education; Eighth edition, 2017.
3. Gerald Keller. Statistics for Management and Economics. Cengage, 2017.
4. Arora, P. N., and Arora, S. CA Foundation Course Statistics. 6th ed. S Chand Publishing, 2018.
5. Mario F Triola. Elementary Statistics. 13th ed., 2018.
6. David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran. Statistics for Business and Economics. 13th Edition, Cengage Learning India Pvt. Ltd., 2019.
7. S D Sharma. Operations Research. Kedar Nath Ram Nath, 2018.
8. Hamdy A. Taha. Operations Research: An Introduction. 10th ed., Pearson, 2016.
9. Malhotra, N. (2012), Marketing Research: An Applied Orientation, 7th ed., Pearson, 2019.
10. Cooper, D.R. and Schindler, P.S. and J. K. Sharma (2018), Business Research Methods, 12th edition, McGraw-Hill Education.

3.10 Self-Assessment Questions

1. Briefly explain the following:

   a. Mutually exclusive events.

   b. Posterior probability.

2. Describe the subjective probabilities of approach.

3. In a group of students, 60% have passed in Economics, 45% have passed in Financial Management and 25% have passed in both. What percentage of students has passed neither?

4. Explain Bayes’ Theorem.

5. Define the following terms: (a) Event, (b) Sample space, (c) Posterior probability.

3.11 Answers to Check Your Progress Questions

1. (d) 35/132

2. (a) 122/147

3. (d) 4/15

4. Degree of Uncertainty

5. (b) If A occurs, B cannot occur

6. Learn more about the event

7. (a) The probability of an event occurring, given that another event has occurred.

8. (c) Classical probability

9. (e) 0.67

10. (a) 1/114

Unit 4

Probability Distribution and Decision Theory

Structure

4.1 Introduction

In the previous unit, you studied concepts of probability, and different types of events. In this unit, you will study Probability distribution, t distribution, F distribution, decision theory, and marginal analysis. Probability distribution is the distribution of all possible outcomes and their probability for any experiment. We can think of Probability distribution as a theoretical frequency distribution. In other words, probability distribution describes the expected variation in outcomes of an experiment. Probability distributions are widely used in making inferences and decisions under condition of uncertainty. It has application in project evaluation, safety stock level, receivables management and other areas of finance.

4.2 Objectives

After going through the unit, you should be able to:

- Recall the concept of random variables;
- Differentiate discrete and continuous probability distribution;
- Explain Binomial, Poisson and hypergeometric distribution;
- Summarize Normal distribution, t distribution and F distribution;
- Evaluate decision-making under conditions of certainty, uncertainty and risk;
- Explain Marginal analysis; and
- State the applications of probability distribution in finance.

4.3 Random Variable

Random variable is a numerical value associated with different outcomes as a result of a random experiment. The term random is used because the outcomes is not known until the experiment has been conducted. Examples of a random variable is given below:

X = Number of vehicles owned by selected family,

X = Number of accidents on a particular place,

X = Number of cars sold during a month by a particular dealer.

Discrete Random Variable

Discrete random variable assumes only the countable numbers for the random variable. For example, the number of accidents in a day cannot be a fraction; similarly, number of cars sold in a month cannot be a fraction, so these are discrete random variables.

Continuous Random Variable

Continuous random variable assumes a number that contains every number within one or more intervals. We cannot count the numbers in continuous random variable as the total numbers between two intervals is infinite. For example, rainfall during a particular period, the height of a person, time taken to reach any place etc. In these examples, rainfall may be 4.5 cm, 21.1 cm or 29 cm. Similarly, height of a person may be 120.5 cm, 122.3 cm, etc.

The following information on variables in COVID diagnosis is an example of random variables.

Example: Variables that Influence COVID Diagnosis and Medication

According to experts of infectious diseases, the following five variables play important roles in influencing the COVID severity. These five variables may vary from person to person.

Microbial dosage: Infection by a number of viral particles

Genetics: This is due to surface proteins on host cells which is like a portal or gateway for viruses

Infection route is the virus' entry path either through mouth or nose.

Virus strength: The spread of virus is due to various degrees of virulent strains.

Immune Status: It is a human being's history of infectious diseases.

These five variables are examples of random variables. Here, Infection route and Immune status are discrete random variables. Remaining three variables – microbial dosage, genetics and virus strength are continuous random variables.

Source: https://publichealth.jhu.edu/2020/five-variables-that-can-affect-covid-19-symptoms-prognosis-and-outcomes (Date: September 18, 2021)

4.4 Probability Distribution of a Discrete Random Variable

Probability of a discrete random variable is a distribution of all discrete random variables and the probability assigned to them. Usually, the past behavior of the variable is studied and the frequency distribution of the past data is ascertained. If the past behavior can be taken as a representative pattern of the future, then the past frequency distribution can be used as a guide for predicting the future values of the random variable. If past behavior cannot help, then subjective probability distribution of the likely future values is formed on the basis of experts' opinion.

Nature/Characteristics of a Discrete Probability Distribution: (i) The probability distribution includes all possible values. (ii) The probability of x is always greater than or equal to 0 and less than or equal to 1. (iii) The values of x are mutually exclusive (only one value can occur in any one experiment). (iv) The sum of the probabilities of each outcome of the experiment adds to one.

Expected Value or Mean of a Discrete Random Variable

Mean of a discrete random variable is also called expected value and is computed by multiplying each random variable by its corresponding probability. The summation gives the expected value of probability distribution. This expected value refers to the value observed on the average. It is the mean value of what we expect to happen if such an experiment were repeated and it does not refer to what would happen in any single random experiment. Mathematically, expected value of probability distribution is expressed as: $E[X] = \sum XP(X)$.

Variance

Variance of probability distribution can be calculated using the following formula: $\sigma^2 = \sum (X - X_e)^2 x\ P_i$ where, X_e = Expected return of distribution,

P_i = Probability of variables.

Standardizing a Random Variable

If X is a random variable with E(X) = μ and V(X) = σ^2, then Y = (X – μ)/ σ is a random variable with mean 0 and standard deviation 1. The standardization (or normalization) of X results in Y.

Covariance

Covariance is a measure of variability between two random variables. The variability of one variable (or data set) (X) in relation to another variable (or data set) (Y) would depend upon three factors: (i) The variability of X which may be measured by the standard deviation σ_x of X. (ii) The variability of Y which may be measured by the standard deviation σ_y of Y. (iii) The correlation between X and Y.

It captures a measure of the correlation of two variables. Positive covariance indicates that as X_1 increases, so does X_2. Negative covariance indicates X_1 decreases as X_2 increases and vice versa. Zero covariance indicates that X_1 and X_2 are uncorrelated. Covariance of variable X with itself is nothing but variance of variance X. In other words, the variance is a special case of covariance.

Computation of Covariance

UNGROUPED DATA

1. **For a population consisting of paired ungrouped data points (X, Y)**

$$\text{Cov}_{XY} = \frac{\sum(X - \mu_X)(Y - \mu_Y)}{N}$$

Where,

μ_x = arithmetic mean of (X),

μ_Y = arithmetic mean of (Y),

N = number of observations in each population.

2. **For a paired sample (X, Y)**

$$\text{Cov}_{XY} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{n - 1}$$

Where,

$\bar{X}$ = arithmetic mean of sample (X),

$\bar{Y}$ = arithmetic mean of sample (Y),

n = number of observations in each sample.

3. **For a Grouped Data**

For grouped data of a paired population

$$\text{Cov}_{XY} = \frac{\sum f(X-\mu_x)(Y-\mu_y)}{\sum f}$$

Where,

f is the frequency of the corresponding (X,Y) values,

μ_X and μ_Y are the means of X and Y.

4. **For a Probability Distribution**

Given a probability distribution of paired data {X, Y} we can compute the covariance using the formula.

$$\text{Cov}_{XY} = \sum[X-E(X)]\,[Y-E(Y)]P(X,Y)$$

Where,

P(X,Y)	=	Joint probability of X and Y,
E(X)	=	Expected value of X,
E(Y)	=	Expected value of Y.

Example 1

Given the following export data for stocks X and Y, calculate the covariance.

Annual Returns (%)

Year	X	Y
1	6.2	–9.5
2	3.6	–11.7
3	4.0	13.8
4	2.4	–5.3
5	0.2	9.5

Covariance of Stocks X and Y

Year	$\left(X_i-\overline{X}\right)$	$\left(Y_i-\overline{Y}\right)$	$\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)$
1	2.92	–8.86	–5.8712
2	0.32	–11.06	–3.5392
3	0.72	14.44	10.3968
4	–0.88	–4.66	4.1008
5	–3.08	10.14	–31.2312
	0	0	–46.144

Covariance:

$$\text{Cov}_{xy} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{n-1} = \frac{-46.144}{4} = -11.536$$

4.5 Discrete Uniform Distribution

Some discrete random variables are of uniform distribution, where all the variables remain constant. For example, in rolling a die the probability of getting 1, 2, 3, 4, 5 or 6 is constant and is represented as 1/6. A random variable with probability distribution is given by:

$P(X = x) = 1/k$ for all values of $x = 0, ..., k$

$P(X = x) = 0$ for other values of x

Where, k is a constant, and is said to follow a uniform distribution.

In fact, $P(X = x) = 1/6$ for all x between 1 and 6. Hence, we have a uniform distribution.

4.5.1 Expectation and Variance

We can find the expectation and variance of the discrete uniform distribution:

Suppose: $P(X = x) = 1/k$ for all values of $x = 0, ..., k$.

Then $E(X) = 1 \text{ x } P(X = 1) + 2.P(X = 2) + ... + k \text{ x } P(X = k)$

$= (1/k) + (2/k) + (3/k) + ... + (k/k)$

$= (1/k)(1 + 2 + ... + k)$

$= (1/k) \text{ x } \frac{1}{2}k\,[2 + (k - 1)]$ (summing the arithmetic progression)

$= \frac{(k+1)}{2}$

It turns out that the variance is*:* $(k^2 - 1)/12$

$$\boxed{\begin{array}{l}\mu = E(X) = (k+1)/2 \\ \sigma^2 = V(X) = (k^2-1)/12\end{array}}$$

4.6 Binomial Distribution

Binomial distribution is a discrete random variable which is used to find the probability of occurrence say, x times out of n trials.

Binomial distribution is the result of a binomial experiment depending on certain conditions. These conditions are summarized below: (i) Binomial experiment consists of a fixed number of trials. (ii) Each trial has only two possible outcomes.

For example, defective or non-defective, success or failure, head or tail, etc. (iii) The probability of the outcome of any trial remains fixed over time. In our example, the probability of the light bulb being defective or non-defective remains fixed throughout. (iv) The trials are statistically independent, that is, the outcome of one trial does not affect the outcome of any other trial. In the above discussed example, the outcome of the light bulb being defective or non-defective does not affect the outcome of any other light bulb being so.

In a binomial experiment, the probability of success (x) in n number of trials is given as:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, ..., n$$

Where,

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

And,

n	=	Number of trials.
x	=	Number of successes in n trials,
p	=	Probability of success,
q	=	Probability of failure = 1 – p.

Example 2

Find the probability of getting exactly 4 heads in 5 tosses of a biased coin, where; P(H) = 3/4 and P(T) = 1/4

$$P(X = 4) = \binom{5}{4} (0.75)^3 (0.25) = 5 \text{ x } (0.75)^3 \text{ x } (0.25) = 0.527$$

for binomial distribution it can be shown as:

μ (mean) = E(x) = np

σ^2 (Variance) = V(X) = npq

4.7 Hypergeometric Distribution

One of the conditions of binomial distribution is that the trials be statistically independent i.e., outcomes of one trial should not effect the outcomes of any other trial, so that the probability of success and failure remains constant. Thus, in an experiment where trials are not statistically independent, the binomial distribution cannot be applied to find the probability of x success in n trials. In this case, we can use the hypergeometric distribution. Normally, such a case arises where the experiment is repeated without replacement. Consider the previous example of bulbs. Suppose Bernoulli's experiment is repeated without replacement. That is, once a light bulb is tested, it is set aside.

If X = number of successes, then:

$$P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, x = 0,1,2,...., n.$$

Thus, for hypergeometric distribution:

$$\boxed{\begin{aligned} &\mu = E(X) = nM/N \\ &\sigma^2 = V(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)} \end{aligned}}$$

It should be clear that if nM/N is small enough, the hypergeometric distribution may be replaced by the binomial distribution. In this unit, we will follow this rule if n/N ≤ 0.05.

Example 3

There are 12 research associates in a research organization. Out of them 7 are females and 5 are males. The organization is planning to select 3 out of the 12 research associates for enrolment into higher education. If 3 research associates are randomly selected out of 12, what is the probability that all 3 of them would be females?

Here,

n = total number of research associates = 12

N = number of selections = 3

M = number of success (female) in the population = 7

N – M = number of failures (male) in the population = 5

x = number of successes in four selections = 3

n – x = number of failures in four selections = 0

$$P(X = 3) = \frac{\binom{7}{3}\binom{5}{0}}{\binom{12}{3}} = 0.1591$$

Thus, probability that all 3 of them would be females is 0.1591.

4.8 Poisson Distribution

Poisson distribution is another discrete distribution which is most commonly used to model the number of random occurrences of certain phenomenon in a specified unit of time or number. It takes into account the value X = 0, 1, 2, 3, ... It can be used to describe various situations, some of which are summarized below: (i) The number of accidents on a road in a particular interval of time. (ii) The number of patients coming to a hospital. (iii) Quality control statistics to count the number of defects of an item. (iv) In insurance problems, to count the number of casualities. (v) In waiting-time problems, to count the number of incoming telephone calls or incoming customers. (vi) The number of traffic arrivals such as trucks at terminals, aeroplanes at airports, ships at docks, and so forth.

Poisson distribution thus, is applicable when counting the number of events in a certain time period or number of objects in a certain volume. Poisson Distribution is represented as:

$$P(r) = \frac{e^{-m} x\ m^{r}}{r!}$$

Where,

r = 0, 1, 2, 3, 4 …

e = 2.7183 (the base of natural Logarithms system)

m = the mean of Poisson distribution i.e., np or the average number of occurrences of an event.

Example 4

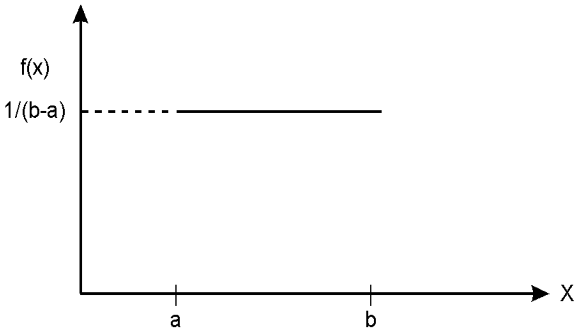
Suppose we want to investigate the safety of a dangerous intersection zone in a particular area. Past hospital records indicate a mean of five accidents per month in this area. Calculate the probability in any month of exactly two accident cases in per month.

$$P(r) = \frac{e^{-m}\ x\ m^{r}}{r!};\ P(2) = \frac{(5)^{2}(e)^{-5}}{2!} = 0.08425$$

4.9 The Continuous Uniform Distribution

In continuous probability distributions, the random variable X can take on any value over its defined range. Consider the height of students in a classroom, it may be 154 cm, 154.525 cm, or 154.64 cm (assume no rounding off). X can take any value from 0 up to 1 but not including 1. The vertical coordinate is a function of x, described as f(x) and referred to as the probability density function and is shown on the Y-axis. The range of possible x values is shown on the horizontal axis. The probability that x will take on a value between a and b on the X-axis is the area under the curve between a and b. The total area under the curve will be equal to 1.

Figure 4.1: The Continuous Uniform Distribution



The probability density function of a typical random variable is given by:

$$f(x)=\begin{cases}1/(b-a), \text{ if} & a \le X \le b\\ 0, & \text{otherwise}\end{cases}$$

Note that the area under the curve equals 1. It can be shown as:

$$\boxed{\begin{array}{l}\mu = E(X) = (a + b)/2\\ \sigma^2 = V(X) = (b - a)^2/12\end{array}}$$

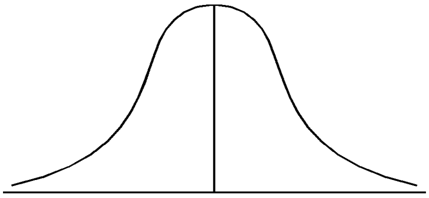
4.10 The Normal Distribution

The normal probability distribution is widely used in continuous probability distribution. The basic reason behind the frequent use of normal distribution is summarized as: (i) Most of the events and populations in the world like the heights and weights of people or the marks of students in a large class tend to follow this distribution. (ii) In normal distribution, observations are found to be clustered around the mean value and their frequency drops sharply as we move away from the mean in either direction. (iii) It can be used to approximate many other discrete and continuous distributions, including the binomial which we studied earlier and many others. If we draw samples of size n (where, n is a fixed number over 30) from any population, then the sample mean $\bar{X}$ will be (approximately) normally distributed with a mean equal to μ i.e., the mean of the population.

The normal distribution when plotted gives a bell-shaped curve such that: (i) The total area under the curve is 1.00. (ii) The curve is symmetric about its center, with an area under each half of the curve of 0.5 units. (iii) The curve has a single peak; thus, it is unimodal. (iv) The mean of a normally distributed population lies at the center of its normal curve. (v) Because of the symmetry of the normal probability distribution, the median and the mode of the distribution are also at the center. (vi) The two tails of the normal probability distribution extend indefinitely and never touch the horizontal axis.

The mean μ and standard deviation σ are the parameters of the normal distribution. The area under a normal distribution curve for any interval can be found if mean μ and standard deviation σ are known. There are a number of normal distributions with different standard deviations and mean.

Figure 4.2: Normal Curve



Area under the Normal Curve

Whatever the value of mean and standard deviation of normal probability distribution, the area under the total curve is always 1. Mathematical research has shown that: (i) Approximately 80 percent of the observation in a normally distributed population lies within ±1.28 standard deviation from the mean. (ii) Approximately 95 percent of the observation in a normally distributed population lies within ±1.96 standard deviation from the mean. (iii) Approximately 98 percent of the observation in a normally distributed population lies within ±2.33 standard deviation from the mean.

Figure 4.3: The Graph of the p.d.f. of a Standard Normal Distribution

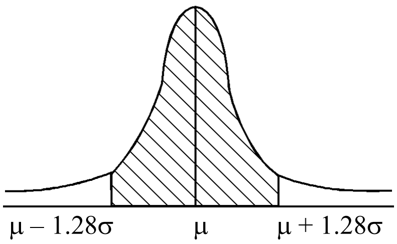


Figure 4.4: The Graph of the p.d.f. of a Standard Normal Distribution

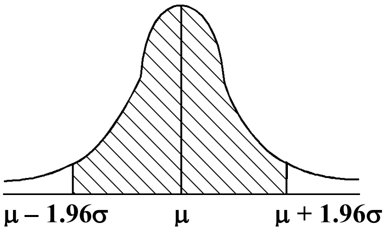
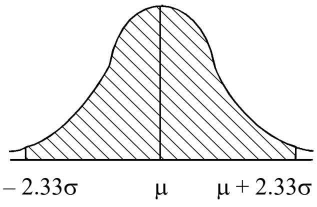


Figure 4.5: The Graph of the p.d.f. of a Standard Normal Distribution



4.10.1 The Standard Normal Probability Distribution

As discussed earlier, there is family of normal distribution curves with different standard deviations and mean. The normal table (in appendix) is used to calculate the area under the curve for different situations. Since it is difficult to provide different tables for different situations we can use standard normal distribution to overcome these problems. The Standard Normal Distribution is a special situation of a normal distribution which has:

Mean (μ) = 0, and Standard Deviation (σ) = 1

4.10.2 Standardizing Normal Variables

In most of the problems, we have mean that is not zero and standard deviations that are not one. Further, normal random variables have different units. In such situations, we convert all normal variables X into a standardized unit Z. Suppose we have a normal population. We can represent it by a normal variable X. Further, we can convert any value of X into a corresponding value Z of the standard normal variable, by using the formula:

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

Where,

X = Value of any random variable,

μ = Mean of the distribution of this random variable,

σ = Standard deviation of this distribution,

Z = Number of standard deviations from X to the mean of this distribution. It is known as the Z score or standard score.

4.11 Lognormal Distribution

If ln(X) is a normally distributed random variable, then X is said to be a lognormal variable. If P1, P2, P3, ... are the prices of a scrip in periods 1, 2, 3, ..., some applications in finance require, (P2/P1), in (P3/P2),... to be normally distributed, that is, continuously compounded returns are required to be normal. This property is described as “Stock Prices Lognormal”. The previous random variables arose out of natural experiments. The following distributions are derived distributions. That is, they are a function of other distributions.

4.12 Normal Approximation to the Binomial

In some situations, we can use normal distribution (continuous) to estimate binomial probabilities (discrete distribution) approximately. The conditions are: (n)(p) > 5 and (n)(q) > 5.

4.13 ‘t’ Distribution

‘t’ distribution, also called student’s ‘t’ distribution, is a theoretical probability distribution and similar to normal distribution in some aspects. t-distribution is symmetrical (bell-shaped) about the mean and it never touches the horizontal axis. But its curve is flatter (lower height and wider spread) than the normal distribution curve and includes some other parameters called degrees of freedom and is centered at zero. The shape of ‘t’ distribution depends on the degrees of freedom, defined as the number of variables that can be chosen freely. Consider the equation X + Y + Z = 10. Once we fix the values of X and Y then the value of Z is set automatically, so the degrees of freedom for this equation is said to be two. Degrees of freedom for ‘t’ distribution is equal to the sample size minus one.

$$\boxed{\text{Degrees of freedom (df) = n – 1}}$$

‘t’ distribution with n – 1 degrees of freedom is defined as:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

Where,

$\overline{X}$ = The sample mean,

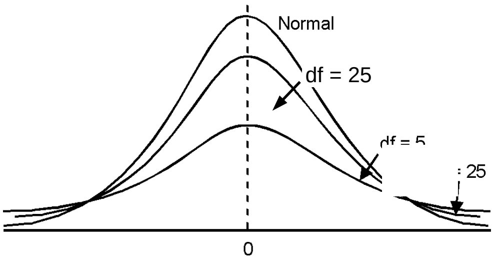
μ = Population mean,

S = Sample Standard Deviation,

n = The sample size.

As shown in the figure 4.6, it is symmetrical like the normal distribution, but its peak is lower than the normal curve and its tail is a little higher above the abscissa than the normal curve.

Figure 4.6: Graph for Normal and ‘t’ Distributions



‘t’ distributions with a smaller degrees of freedom has more area in the tail of the distribution than one with a larger degrees of freedom. As the degrees of freedom for a ‘t’ distribution gets larger and larger, ‘t’ distribution gets closer and closer to the standard normal distribution. As the df increases, ‘t’ distribution approaches the standard normal distribution. The standard normal curve is a special case of ‘t’ distribution when df = ∞.

4.13.1 Use of ‘t’ Distribution

The most important use of ‘t’ distribution is in calculating the confidence interval (discussed later). As discussed earlier distribution approaches, the standard normal distribution when degrees of freedom approaches ∞. But in some practical case, when the sample size increases to 30, the results for t distribution and normal distribution are approximately equal. So, the best use of ‘t’ distribution is when the degrees of freedom is less than 30. Another condition for using ‘t’ distribution is when the population standard deviation is unknown.

$n \leq 30$	– population standard deviation is unknown	– Use t distribution
$n > 30$	– Population standard deviation is known	– Use Z distribution

Consider ‘t’ distribution with df = 13. What is the area to the right of 1.771? From the ‘t’ distribution table given (in appendix) it can be seen that the area is 0.05.

4.14 F Distribution

The shape of F-distribution also depends on the degrees of freedom. However, F-distribution uses: degrees of freedom for the numerator and degrees of freedom for the denominator. n_1 denotes the degrees of freedom for numerator and n_2 for denominators. Each combination of degrees of the freedom will give a different shape of the curve. Both random variables have yet another distribution, called the χ^2 distribution.

Example 5

Consider an F distribution with degrees of freedom 2 in the numerator and 13 in the denominator. What is the area under the curve, right of 3.81?

From the F distribution table given (in appendix), this area equals 0.05.

Check Your Progress - 1

1. Find the mean and variance of a binomial distribution with n = 12 and p = 0.45.

   a. $\mu = 5.40, \sigma2 = 2.97$

   b. $\mu = 2.97, \sigma2 = 5.4$

   c. $\mu = 5.40, \sigma2 = 1.72$

   d. $\mu = 2.43, \sigma2 = 2.97$

   e. $\mu = 5.40, \sigma2 = 2.43$.

2. The board of ABC company is considering some major policy changes that will affect the company’s three divisions. The management was keen on getting the employees’ opinion before going ahead with the changes.

Meetings were held with groups of employees from the three divisions. The opinions of the employees are summarized in the table given below.

Opinion	Textile	Garments	Finance
Strongly oppose	2	2	4
Slightly oppose	2	4	3
Neutral	3	3	5
Slightly support	2	3	2
Strongly support	6	3	1

An employee from one of the groups strongly opposes the policy changes.

What is the probability that he is from the Finance division?

a. 4/45.

b. 1/4.

c. 4/15.

d. 1/2.

e. 3/4.

3. An investor has collected data of a company's share prices over 30 days and arrived at the figures that it is uniformly distributed price between 465 and 635. What is the mean and variance of this price variable?

a. 1350,49

b. 2700,2408

c. 1350,2700

d. 550,2408

e. 1000,500

4. Which distribution is used in continuous probability distribution?

a. Binomial distribution

b. Hypergeometric distribution

c. Continuous uniform distribution

d. Normal distribution

e. Poisson distribution

5. Suppose a population consists of 100 items, 54 of which are successes. IF a random sample drawn from that population consists of 20 items, what is the probability that, 11 of these are successes?

a. 196

b . 168

c. 212

d. 180

e. 200

6. Which of the following is similar to normal distribution in some aspects?
   a. t distribution
   b. Binomial distribution
   c. Poisson distribution
   d. F-distribution
   e. Hypergeometric distribution

4.15 Decision Theory

Decision theory is commonly utilized in daily life, perhaps without even analyzing its contents. It focuses on only some aspects of human activity, in particular, how we use our freedom. There are different situations, treated by decision theorists, in which we have options to choose from. In these situations, our choices reflect goal-directed behavior in the presence of options. The primary purpose of this analysis is to increase the likelihood of a good outcome by making a good decision. Good decision means a decision that is consistent with the information and preference of the decision maker. Thus, decision analysis attempts to provide a decision-making framework based on any and all information, whether in terms of sample information, judgment information or both.

Our discussion of decision theory is based on the following assumptions: (a) The decision maker can define all decision alternatives or strategies which are being considered. (b) He can define the various states of nature for the decision setting which are not under his control. (c) He can estimate quantitatively the consequences (benefits or costs) of selecting any decision alternative, having any state of nature occurrence. Based on the level of knowledge, decision analysis can be divided into three major categories:

i. **Decision-making under the Conditions of Certainty:** Decision makers are completely certain about the state of nature which is going to occur.

ii. **Decision-making under the Conditions of Uncertainty:** Decision makers have no knowledge about the likelihood of occurrence of the various states of nature.

iii. **Decision-making under the Conditions of Risk:** Decision makers have sufficient knowledge about the states of nature to assign probabilities to the likelihood of their occurrence.

4.16 Decision-making under Conditions of Certainty

Usually, the alternatives and their outcomes are known in advance. A Finance Manager often faces the choice of whether to invest in growth stock, bond, commodity market, money market, etc. Normally, various possible scenarios are available. The decision maker identifies many options and effectively evaluates them. Conditions of certainty tend to be rare, especially when significant decisions are involved.

Example 6

A Finance Manager must make a decision about investing in one of the three investment portfolio packages. Each investment portfolio contains different proportions of growth stock, bonds and real estate. The table below gives the gains of investing in these portfolios under different states of nature based on the projection of changes in stock prices, yields on bonds and appreciation in real estate values.

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)
A	2,50,000	32,000	– 21,000
B	30,000	37,000	6,000
C	3,50,000	45,000	– 4,000

Under conditions of certainty, decision-making becomes very easy. The Finance Manager can choose that alternative which gives him the maximum profit. Thus, the analyst decides to invest in portfolio C in the stagnant period, in slow growth period and opts for portfolio B in rapid growth because that alternative gives as well as him the maximum benefit.

4.17 Decision-making under Conditions of Uncertainty

Under uncertain circumstances, the decision maker is not certain about the state of nature and its probability. As these are unknown, the decision maker can take the decision based on different pay-offs. Depending on the decision maker’s outlook (whether optimistic or pessimistic) different approaches can be used.

4.17.1 Maximax

Maximax is an optimistic approach where the decision maker believes that, given the selection of any strategy, “nature” will act in a way which provides the greatest reward. Thus, the decision maker selects the maximum profit associated under each alternative and then selects the alternative that produces the highest of these maximum pay-offs.

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)	Maximum of each Alternative (Rs.)
A	2,50,000	32,000	–21,000	2,50,000
B	30,000	37,000	6,000	37,000
$\boxed{\text{C}}$	3,50,000	45,000	–4,000	3,50,000

The maximum of each alternative as shown in the above table will give us the alternative to be chosen. Thus, in the above example, the maximum benefit is derived from the alternative C. The analyst believes that the nature will behave as per the stagnant period.

4.17.2 Maximin Criterion

It is a pessimistic approach, wherein the decision maker assumes that the worst will happen. The approach here is to minimize the losses. Thus, the decision maker first selects the minimum profit associated under each alternative and then selects the alternative that produces the highest of this minimum pay-off.

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)	Minimum of Each Alternative (Rs.)
A	2,50,000	32,000	–21,000	–21,000
B	30,0000	37,000	6,000	6,000
C	3,50,000	45,000	–4,000	–4,000

Continuing with the earlier example, the decision maker will choose maximum of $\{-21,000,\ \ 6,000,\ \ -4,000\}$ *i.e.* $6,000.$ Thus, in the above example, the decision maker chooses to invest in portfolio B. This approach guarantees the minimum profit.

4.17.3 Hurwicz Criterion

Hurwicz criteria is applied where decision makers are neither optimistic nor pessimistic. Their outlook lies in between these two. In this criteria, decision maker specifies his level of optimism by an index of optimism, α which is assigned a value between 0 and 1, with both the numbers being inclusive. The decision is made based upon weighted profits. Weighted profits are calculated as:

α (maximum profit for alternative) + $(1 - \alpha)$
(Minimum profit for alternative).

The alternative which gives maximum of the weighted profits is the decision maker's choice. This approach is known as the **Hurwicz Criterion.** Assume that the decision maker specifies the value of α as 0.7. He has a tendency towards optimism.

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)	Strategy: Investment in Portfolio (Rs.)
A	2,50,000	32,000	–21,000	A
B	3,00,000	37,000	6,000	B
C	3,50,000	45,000	–4,000	C

The weighted profits are calculated as follows:

Strategy	Weighted Gains	Rs.
A	0.7 (250,000) + 0.3 (–21,000)	1,68,700
B	0.7 (3,00,000) + 0.3 (6,000)	2,11,800
$\boxed{C}$	0.7 (3,50,000) + 0.3 (–40,00)	2,43,800

Thus, the best alternative is to invest in portfolio C.

4.17.4 Regret Criterion

Regret criterion is based upon the opportunity loss or opportunity cost. Opportunity loss means amount of pay-off that has been incurred by not selecting the best alternative. The regret is measured by the difference between the maximum profit that we would have realized in case of known state of nature and the profit we actually realize. For example, if the decision maker wants to calculate the regret for minimax he/she can prepare the regret table as shown below.

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)
A	3,50,000 – 2,50,000	45,000 – 32,000	6,000 + 21,000
B	3,50,000 – 30,0000	45,000 – 37,000	6,000 – 6,000
C	3,50,000 – 3,50,000	45,000 – 45,000	6,000 + 4,000

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)	Maximum Regret (Rs.)
A	1,00,000	15,000	27,000	1,00,000
B	50,000	8,000	0	50,000
C	0	0	10,000	10,000

From the table, we can say that if state S_1 occurs and alternative C is chosen we would have no regret. Now we have to find the maximum regret for each alternative as shown in the table. The alternative which gives us the minimum of these maximum regrets will be the decision. Thus, as far as this criterion is concerned, it is best to invest in alternative C. It is also designed for the pessimist and assumes that the nature will respond, so as to maximize the level of regret for any alternative chosen. Similarly, we can draw the regret criteria for every alternative.

4.18 Decision-making under Conditions of Risk

In decision-making under conditions of risk, all possible states of nature and their respective probabilities are known but decision makers are uncertain about which state of nature will occur. Probability is normally assigned to the different states of nature based either on the decision maker's feelings and experience or on the collection and analysis of numerous data related to the states of nature. Using these probabilities, the decision maker can develop his strategies. Followings are some of the popular methods in use: (i) Expected Monetary Value (EMV). (ii) Expected Value of Perfect Information (EVPI). (iii) Decision Tree Analysis.

4.18.1 Expected Monetary Value (EMV)

Expected monetary value is calculated by multiplying the probabilities of each state of nature by the state's associated pay-off. Finally, the result for each decision alternative is summed up and the alternative which gives the highest expected value is selected.

Example 7

A fruit merchant purchases apples from wholesalers for Rs.200 each box. The merchant will initially sell the box for Rs.300 each. He will then sell all the unsold boxes for Rs.100 each to another fruit merchant. Historical data confirms that daily demand for sales for the item assumes four possible values. The table gives the demand information along with the respective probabilities of occurrence. The merchant is trying to decide how many units of the item to stock in a month. Its goal is to select the quantity which maximizes expected monthly profit.

Estimate of Demand	Probability
3	0.20
4	0.25
5	0.45
6	0.10

To calculate the optimum stock level which maximizes the profit we have to construct the conditional profit table. This table summarizes the daily profit which would result in the selection of a particular stock level and the occurrence of a specific level of demand. The table also reflects the losses that occur when the remaining stock is sold to the discounter at the end of the day and it does not take into account the additional profit it lost when customers demand more than the store has stocked.

Conditional Profit Table

Stock Decision (Probability)	Possible Demand			
	3 (0.2)	4 (0.25)	5 (0.45)	6 (0.1)
3	3,00	3,00	3,00	3,00
4	200	400	400	400
5	100	300	500	500
6	0	200	400	600

The conditional profit values are determined by computing the total profit from units sold and subtracting from this any loss which would have to be absorbed because of overstocking. For example, if the fruit merchant decides to stock 3 boxes it always results in a conditional profit of Rs.300 [3 (300 – 200)], because if the demand is for more than 300 boxes, all the boxes stocked will be sold. But, if he decides to stock 4 boxes and the demand is for 3 boxes, the conditional profit is equal to the total profit of selling 3 boxes and the loss incurred by overstocking 1 box, where loss incurred because of overstocking is the difference between the cost of the remaining boxes and the sale price of these boxes to another fruit merchant for Rs.100. Thus,

Conditional profit = (3 x 100) – (1 x 100) = Rs.200

The conditional profits for other stock decisions are also calculated in a similar manner. The expected daily profits for each stock decision can be determined by weighing each conditional profit by its likelihood of occurrence (which is the probability of the corresponding level of demand). The table below gives the conditional profit for each level of stock decision.

Expected Monthly Profit Computation

Stock 3 boxes

Conditional Profit (Rs.)	Probability of Occurrence	Expected Daily Profit (Rs.)
300	0.20	60
300	0.25	75
300	0.45	135
300	0.10	30
		300

Stock 4 boxes

Conditional Profit (Rs.)	Probability of Occurrence	Expected Daily Profit (Rs.)
200	0.20	40
400	0.25	100
400	0.45	180
400	0.10	40
		360

Stock 5 boxes

Conditional Profit (Rs.)	Probability of Occurrence	Expected Daily Profit (Rs.)
100	0.20	20
300	0.25	75
500	0.45	225
500	0.10	50
		370

Stock 6 boxes

Conditional Profit (Rs.)	Probability of Occurrence	Expected Daily Profit (Rs.)
0	0.20	0
200	0.25	50
400	0.45	180
600	0.10	60
		290

On the basis of expected monthly profits, the best decision is to stock 5 boxes, resulting in an expected (average) profit of Rs.370.

Example 8

The Portfolio Manager of an Asset Management Company wants to take a decision regarding investment in a portfolio. He has collected the information on different portfolios and their profit in different states of nature.

The information about different states of nature and their portfolio is given below:

Strategy: Investment in Portfolio	Stagnant (Rs.)	Slow Growth (Rs.)	Rapid Growth (Rs.)
	P = 0.4	P = 0.25	P = 0.35
X	25,000	32,000	(21,000)
Y	30,000	37,000	6,000
Z	35,000	45,000	(4,000)

We can evaluate the decision of the Portfolio Manager relating to the investment in portfolios X, Y and Z.

Calculation of Expected Gains

Portfolio X

Gains	Probability	Expected Gains
25,000	0.4	10,000
32,000	0.25	8,000
–21,000	0.35	(7350)
		10,650

Portfolio Y

Gains	Probability	Expected Gains
30,000	0.4	12,000
37,000	0.25	9,250
6,000	0.35	2,100
		23,650

Portfolio Z

Gains	Probability	Expected Gains
35,000	0.4	14,000
45,000	0.25	11,250
(4,000)	0.35	–1,400
		23,850

Thus, as per the expected value criteria the investment analyst should decide to invest in portfolio Z given the probabilities of 0.4, 0.25 and 0.35 for Stagnant, slow growth and rapid growth respectively.

4.18.2 Expected Value of Perfect Information

The expected value of perfect information is the difference between the expected profit under the situation where the decision maker knows which state of nature would occur and the best expected monthly profit when there is no information about the occurrence of the state of nature. i.e.,

EVPI = Expected monetary pay off with perfect information – Expected monetary value without information

In the previous example, we can calculate the expected profit under perfect information, assuming that the uncertainty has been removed by using the conditional profit table as shown below.

Conditional Profit Table under Perfect Information

Stock Decision (Probability)	Possible Demand			
	3 (0.2)	4 (0.25)	5 (0.45)	6 (0.1)
3	300	–	–	–
4	–	400	–	–
5	–	–	500	–
6	–	–	–	600

If the fruit merchant estimates future demand to be 3 boxes, he stocks only 3 boxes and makes a profit of Rs.300. The profit values for other levels of stock are calculated similarly.

Thus, with perfect information the merchant can realize profits as under:

Expected Profit under Certainty

Stock Decision	Conditional Profit	Probability	Expected Profit Under Certainty
3	300	0.20	60
4	400	0.25	100
5	500	0.45	225
6	600	0.10	60
			445

The expected profit under certainty is Rs.445. Thus, the maximum possible expected profit is Rs.445.

Thus, in the above problem, expected value of perfect information is Rs.75(445 – 370).

4.18.3 Decision Tree Analysis

Decision trees are excellent mathematical tools that can help the decision maker to choose between several courses of action. They provide a highly effective structure within which he can layout various options and consider their possible outcomes. They can also help in forming a balanced picture of the risks and rewards associated with each possible course of action.

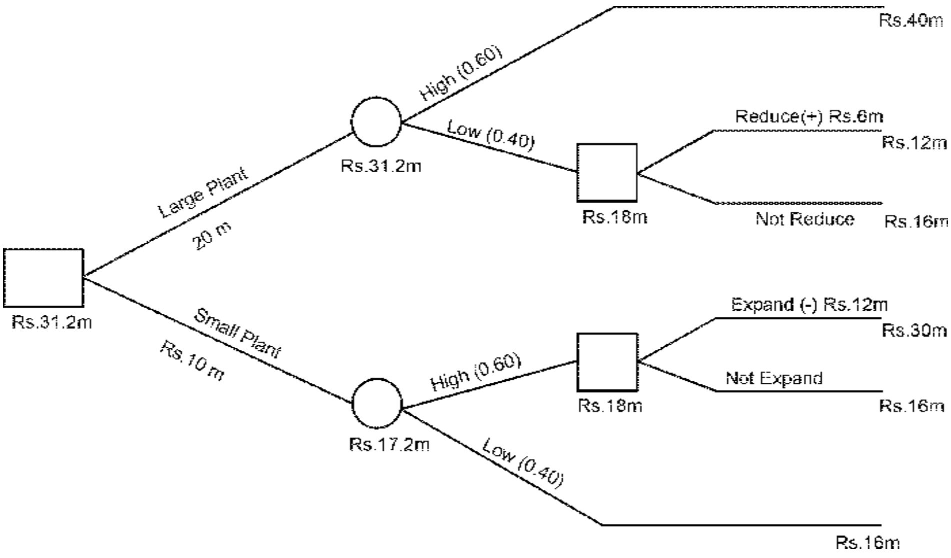
For example, the Finance Manager of 'Johnson's Baby Shampoo,' with successful first two years of the company's operations, is considering setting up another plant to meet the growing demand for the product. If the demand is very high the new project is expected to earn Rs.10 million per annum and if the demand is low it is expected to earn Rs.4 million per annum. The probability that the demand is very high is 0.60.

The firm can set-up either a large plant costing Rs.20 million or a small plant costing Rs.10 million. Either plant will take one year to build and as such no profits are earned in year 1. Both the plants have an estimated life of 4 years after which their value would be considered negligible. In case of high demand, the firm can expand the small plant in year 2 at an additional cost of Rs.12 million. Alternatively, in case of low demand, the firm can reduce the large plant in year 2 and recover Rs.6 million. Expansion or reduction of plant size will take a whole year, so that no profits will be earned in year 2.

Steps to Draw a Decision Tree

The decision tree can be drawn using the following steps: (i) Draw a small square (called decision node from which one of several alternatives may be chosen) towards the left of a large piece of paper. (ii) From this box draw out lines towards the right for each possible solution, and write that solution along the line. Keep the lines apart as far as possible so that you can expand your thoughts. (iii) At the end of each line, consider the results. If the result of taking that decision is uncertain, draw a small circle. (Known as the state of nature node out of which one state of nature will occur.). (iv) If the result is another decision that you need to make, draw another square. Squares represent decisions, and circles represent uncertain outcomes. Write the decision or factor above the square or circle. If you have completed the solution at the end of the line, just leave it blank. The decision tree for the above discussed example is given below:

Figure 4.7: Decision Tree



How to Solve?

After drawing the tree, the Finance Manager can evaluate the decision tree from right to left. At every decision node, the decision maker chooses that alternative which gives him the highest worth. Next, at each circle, the probability of each outcome can be written. Now the value at the state of nature node is obtained by adding the product of the incomes of the various states of nature and their respective probabilities. In the above example, if a large plant is set-up and demand is reduced, the income from the project would be Rs.18 million [i.e. (4 x 3) + 6]. In case its size is not reduced, income would be only Rs.12 million. Hence, if a large plant is set-up and its size is reduced, the expected value would be Rs.31.02 million. If a small plant is set-up and high demand is expected, income would be Rs.18 million, if not the income would be Rs.12 million. Thus, if a small plant is set-up and its size is expanded the total expected value from the project would be Rs.17.02 million. Finally, the Finance Manager can select to set-up a large plant and in case of low demand reduce its size as the profit is maximum in this case.

4.19 Marginal Analysis

In the earlier example of optimum stock level if the number of scenarios and possible actions are too large, then it is very difficult to calculate the optimum level of stock using the conditional table. The marginal analysis avoids the unnecessary calculation. Marginal analysis is based on the fact that the additional unit bought can be sold or unsold.

For any additional unit bought, if the probability of sold unit is p then probability of unsold unit is (1 – p). If we know the probability of additional units sold, we can easily calculate the increase in the conditional profit by selling an additional unit. The unit profit is known as marginal profit. In other words:

> Marginal Profit (MP) is the additional profit generated by increasing our activity level by one unit.

Similarly, marginal loss is nothing but loss incurred by increasing our activity level by one unit. Let us consider the earlier example of optimal stock, the conditional profit is Rs.30 for a stock of 3 boxes, and the marginal profit in selling one unit is Rs.100 (300 – 200) whereas marginal loss is Rs.100 (200 – 100) i.e., if unit product is sold, the conditional profit will increase by Rs.100, but if the additional product is not sold then the conditional profit will be reduced by Rs.100 i.e., we can increase additional stock till the expected marginal profit is greater than expected marginal loss. Now we can calculate the probability of selling additional unit at every stock level.

The probability of different demand levels is given as follows:

Estimate of Demand	Probability
3	0.20
4	0.25
5	0.45
6	0.10

From the above table, we can say the probability of selling more than 3 units is (1 – 0.20) i.e., 0.80. Similarly, probability of selling more than 5 boxes is 0.75.

The break even level at which marginal loss for selling and stocking additional unit is equal to the marginal profit for selling and stocking one additional unit will be: P x MP = (1 – P) ML:

Solving this equation, we can get the minimum probability at which marginal profit equals to marginal loss as: $P^* = ML/(ML + MP)$.

Now we can check the MP and ML at each Stock Level

Stock Level = 3		
MP = 100	P = P(D > 3) = 0.80	E(MP) = 0.8 x 100 = 80
ML = 100	(1 – P) = P (D ≤ 3) = 0.20	E(ML) = 0.20 x 100 = 20
Stock Level = 4		
MP = 100	P = P(D > 4) = 0.55	E(MP) = 0.55 x 100 = 55
ML = 100	(1 – P) = P(D ≤ 4) = 0.45	E(ML) = 0.45 x 100 = 45
Stock Level = 5		
MP = 100	P = P(D > 5) = 0.10	E(MP) = 0.10 x 100 = 10
ML = 100	(1 – P) = P(D ≤ 5) = .90	E(ML) = 0.9 x 100 = 90
Stock Level = 6		
MP = 100	P = P(D > 6) = 0	E(MP) = 0.90 x 100 = 90
ML = 50	(1 – P) = P(D ≤ 6) = 1	E(ML) = 0.10 x 100 = 10

In the above table, the expected marginal profit goes on decreasing and the expected marginal loss goes on increasing. Therefore, at some level,

Expected (MP) = Expected (ML).

That is, P(MP) = (1 – P)ML

Solving for P, we have, $P^* = ML/(ML + MP)$

(In our example, P* represents the minimum required probability of selling at least one additional unit to justify the stocking of that additional unit.)

$$P^* = \frac{100}{100+100} = \frac{100}{200} = 0.50$$

At 4 units probability is 0.55 which is greater than 0.50 i.e., $P > P^*$, which denotes that, we can go ahead for one more stock and that a stock of 5 units would be the optimum level. Marginal analysis may be applied to continuous data too.

Example 9

A Chinese Restaurant prepares a chicken dish, 'Chicken Fun', every Sunday. The Restaurant Manager wants to ensure that the restaurant earns the maximum profit on this dish. A single whole chicken goes to prepare, two portions of Chicken Fun. The selling price per portion of the dish is Rs.50, which is considered to be a bargain price, which has made the dish very popular. Each portion of the dish costs Rs.32 including labor and other costs of preparation. It was observed over a long period of time that the demand for the dish was normally distributed with a mean of 210 portions and a standard deviation of 50 portions. The dish is perishable as the unsold portions cannot be preserved and sold later.

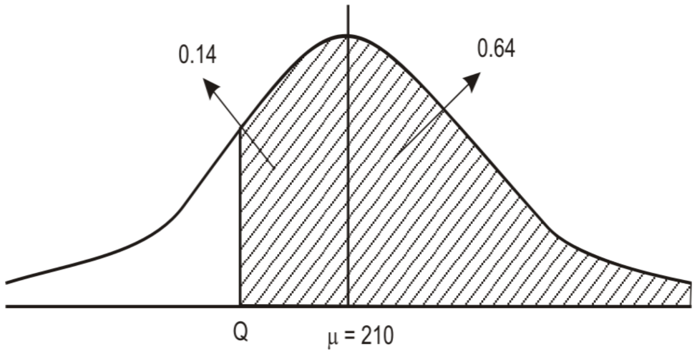
How many whole chickens should be ordered by the Restaurant Manager?

Minimum required probability of selling each portion of the dish,

Marginal profit, MP = 50 – 32 = Rs.18

Marginal loss, ML = Rs.32

$$p^* = \frac{\text{ML}}{\text{MP}+\text{ML}} = \frac{32}{32+18} = 0.64$$



The Z-value for the area 0.14 to the left of the mean is –0.36.

$\therefore -0.36 = \frac{x-210}{50}$; x = 210 – (0.36 × 50) = 192

∴ The restaurant should prepare 192 portions at the optimal level.

∴ The number of chickens that should be ordered = $\frac{192}{2}$ = 96.

4.20 Applications in Finance

The word probable roughly means "likely to occur" in the case of possible future occurrences, or "likely to be true" in the case of inferences from evidence. The application of probability theory is very common in today's world. Some of the areas where probability and probability distribution are widely used are discussed below:

Project Evaluation

Effective value criteria are often used to evaluate a project. Market researchers can predict the probability under different states of nature and using the decision tree analysis they can decide and calculate the NPV of a project to decide whether to accept it or not. The decision can be taken at a particular node also, where managers have to decide whether they need to expand.

Safety Stock Level

Probability method is sometimes used to calculate the level of safety stock. As per the EOQ (Economic Order Quantity) model:

Reorder level = Lead time in number of days x Usage in units per day

EOQ model assumes that the usage is uniform and that the lead time can be predetermined and does not tend to vary. But, in practice, the used in units per day may not be predictable and may move in a random manner and similarly the lead time could also be a random variable. This could lead to stock out and hence potential losses to the firm. One way to reduce stock out loss is to maintain safety stock. But the level of safety stock should be at the optimal level at which the stock-out costs and carrying costs are minimized. By estimating probability distribution of usage during the reorder period, we can easily calculate the optimal safety stock level.

Receivables Management

Probability is frequently used in receivables management to decide whether or not to grant credit to a particular customer. Usually, a certain subjective probability of the payment pattern of the customer can be specified based on an analysis of the customer's financial viability, the nature of profession or employment or business he is engaged in, his relationship with other suppliers, his past record of payments to other suppliers, etc. The decision to grant credit or not is usually based on expected value calculations.

Check Your Progress - 2

7. Which of the following conditions is considered while decision making by a finance manager who is keen on making profit under different investment situations?
   a. Risk
   b. Uncertainty
   c. Certainty
   d. Maximax
   e. Maximini
8. Which of the following is not a method used for decision making under conditions of risk?
   a. Decision tree analysis
   b. Expected monetary Value

c. Expected Value of Perfect Information

d. Marginal Analysis

e. Expected gains

9. A Finance Manager has to make a decision about investing in one of the four investment portfolio packages. The table below gives the gains of investing in these portfolios. Using Maximax approach, which portfolio she will choose?

Portfolio	Stagnant growth	Slow growth	Rapid growth
A	25000	32000	33000
B	30000	37000	41000
C	35000	47000	47000
D	44000	45000	46000

a. A

b. B

c. C

d. D

e. None of the above

10. When the decision maker possesses information about the probabilities of the possible states of nature, his/her decisions are said to be made under conditions of

a. Uncertainty

b. Risk

c. Certainty

d. Probability

e. None of the above

4.21 Summary

- Probability distributions are most commonly used in decision-making and can be considered as theoretical frequency distributions. In discrete probability distribution only whole numbers are taken as a variable. In continuous probability distribution, a variable can be considered as any value within a certain range. Binomial and Poisson Distribution are discrete probability distributions. Normal probability distributions are continuous probability distributions.

- Decision-making can be under conditions of certainty, under conditions of uncertainty and under risk. When the decision maker is aware of all the possible states of nature and has sufficient information to assign any probabilities of occurrence to the various states of nature, the appropriate criterion of decision-making is the expected value criterion. Under uncertain conditions, a decision maker may use any of the given criteria depending on his nature. For example, the maximax criterion applies to an optimistic person, maximin to a pessimistic person, the Hurwicz criterion to subjective judgment and regret criterion to relevant opportunity cost. Finally, probability distribution can be used for project evaluation, finding safety stock levels and receivables management.

4.22 Glossary

Bernoulli's Process: A process in which each trial has only two possible outcomes. The probability of the outcome of any trial remains fixed over time, and the trials are statistically independent.

Binomial Distribution: It is discrete random variable which is used for finding the probability of occurrence (x times) out of n trials.

Continuous Random Variables: Have uncountably many possible values, and take each with probability 0; these quantities usually represent lengths, weights, etc., and need not be integers.

Covariance: Covariance is a measure of variability between two random variables.

Discrete Random Variables: It can only take a finite or countable number of values, and have a positive probability of taking each one; typically these are integer-valued quantities obtained by counting.

Expected Value: A weighted average of the outcomes of an experiment.

Poisson Distribution: A discrete random variable which is used to count the number of events that occur in a certain time interval or spatial area.

Probability Distribution: The probability distribution of a random variable is a list of probabilities associated with each of its possible values.

Marginal Profit/Loss: The profit/loss incurred from stocking a unit that is not sold.

Normal Distribution: A distribution of a continuous random variable with a single-peaked, bell-shaped curve. The mean lies at the center of the distribution, and the curve is symmetrical around a vertical line erected at the mean. The two tails extend indefinitely, never touching the horizontal axis.

4.23 Suggested Readings/Reference Material

1. Gupta, S. P. Statistical Methods. 46th Revised ed. New Delhi: Sultan Chand & Sons. 2021.
2. I. Levin Richard, H. Siddiqui Masood, S. Rubin David, Rastogi Sanjay. Statistics for Management. Pearson Education; Eighth edition, 2017.
3. Gerald Keller. Statistics for Management and Economics. Cengage, 2017.
4. Arora, P. N., and Arora, S. CA Foundation Course Statistics. 6th ed. S Chand Publishing, 2018.

5. Mario F Triola. Elementary Statistics. 13th ed., 2018.
6. David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran. Statistics for Business and Economics. 13th Edition, Cengage Learning India Pvt. Ltd., 2019.
7. S D Sharma. Operations Research. KEDAR NATH RAM NATH, 2018.
8. Hamdy A. Taha. Operations Research: An Introduction. 10th ed., Pearson, 2016.
9. Malhotra, N. (2012), Marketing Research: An Applied Orientation, 7th ed., Pearson, 2019.
10. Cooper, D.R. and Schindler, P.S. and J. K. Sharma (2018), Business Research Methods, 12th edition, McGraw-Hill Education.

4.24 Self-Assessment Questions

1. Explain the concept of t distribution. In which situation can we use the t distribution?

2. Poisson distribution is applicable when an event occurs at random points in time or space. Discuss briefly Poisson distribution and point out its role.

3. Decision tree is considered as a mathematical model of the decision situation and is a useful tool for decision-making for manager. It requires the decision maker to examine all possible outcomes, desirable and undesirable, their chance of occurring, etc. Discuss the various advantages of using decision trees for problem-solving.

4. Describe the concept of marginal analysis in decision-making.

5. Write short notes on (a) Maximax criterion, (b) Hurwicz criterion.

4.25 Answers to Check Your Progress Questions

1. (a) 5.4, 2.97

2. (d) 1/2

3. (d) Mean:550, Variance:2408

4. Normal distribution

5. (a) 196

6. (a) t distribution

7. (c) Certainty

8. (a) Decision tree analysis

9. (c) C

10. (c) Certainty

Quantitative Methods

Course Structure

Block	Unit Nos.	Unit Title
I Introduction to Statistics and Probability		
	1.	Arranging Data
	2.	Central Tendency and Dispersion
	3.	Probability
	4.	Probability Distribution and Decision Theory
II Statistical Relations and Hypothesis Testing		
	5.	Statistical Inference and Hypothesis Testing
	6.	Correlation and Linear Regression
III Statistical Regression and Quality Control		
	7.	Multiple Regression
	8.	Time Series Analysis
	9.	Quality Control
IV Statistical Distributions, Variations and IT		
	10.	Chi-Square Test and Analysis of Variance
	11.	Role of IT in Modern Business Enterprise
	12.	Statistical Software Tools
V Advanced Statistics		
	13.	Index Numbers
	14.	Simulation
	15.	Linear Programming
VI Business Research		
	16.	Introduction to Business Research Methods
	17.	Questionnaire Design
	18.	Report Writing